



ALÉM DAS NUVENS:
EXPANDINDO AS FRONTEIRAS DA CIÊNCIA DA
INFORMAÇÃO

Isa M. Freire, Lilian M. A. R. Alvares,
Renata M. A. Baracho, Mauricio B. Almeida,
Beatriz V. Cendon, Benildes C. M. S. Maculan
(Org.)



ALÉM DAS NUUVENS:
EXPANDINDO AS FRONTEIRAS DA CIÊNCIA DA
INFORMAÇÃO

ISSN 2177-3688

BELO HORIZONTE

ECI/UFMG

2014

COMISSÃO CIENTÍFICA

Profa. Dra. Renata Maria Abrantes Baracho – UFMG: Presidente
Profa. Dra. Lillian Alvares – UnB
Profa. Dra. Icléia Thiesen – Unirio
Profa. Dra. Brígida Maria Nogueira Cervantes – UEL
Profa. Dra. Giulia Crippa - USP
Profa. Dra. Emeide Nóbrega Duarte – UFPB
Prof. Dr. Clóvis Montenegro de Lima – IBICT
Profa. Dra. Aida Varela - UFBA
Profa. Dra. Leilah Santiago Bufrem – UFPE
Profa. Dra. Plácida Amorim da Costa Santos – Unesp/Marília
Profa. Dra. Luisa M. G. de Mattos Rocha – IPJB/RJ
Prof. Dr. Carlos Xavier de Azevedo Netto – UFPB
Profa. Dra. Maria Cristina Soares Guimarães - IBICT/Fiocruz

PARECERISTAS DA COMISSÃO CIENTÍFICA DO GT 2

ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO

Ana Célia Rodrigues
Ana Cristina de Albuquerque
Brígida Maria Nogueira Cervantes
Carlos Henrique Marcondes de Almeida
Célia da Consolação Dias
Cristina Dotta Ortega
Dulce Amélia de Brito Neves
Evelyn Goyannes Dill Orrico
Fabio Assis Pinho
Fernanda Passini Moreno
Gercina Ângela Borem de Oliveira Lima
João Batista Ernesto de Moraes
José Augusto Chaves Guimarães
Lígia Café
Lilian Alvares
Marcos Luiz Cavalcanti Miranda
Maria Elizabeth Baltar Carneiro de Albuquerque
Mariângela Spotti Lopes Fujita
Marisa Bräscher
Maria Luiza de Almeida Campos
Maurício Barcellos Almeida
Miriam Paula Manini
Nair Yumiko Kobashi
Renato Rocha Souza
Renato Tarciso Barbosa de Sousa
Rosa Inês de Novais Cordeiro
Rosali Fernandez de Souza
Rosane S. A. Lunardelli
Vânia Mara Alves Lima
Vera Dodebei
Walter Moreira

Realização



Agências de Fomento



Grupos de pesquisa



Grupo de Pesquisa e Estudos em Museologia
Arte e Ciência na Tecnologia, Educação e Ciências



CIÊNCIA DA INFORMAÇÃO - UFMG



Apoio



UNIVERSIDADE
FUMEC
DE MINAS GERAIS PARA O MUNDO

De olho na CI



NITEG
UFMG



A água de Minas



GT 2

ORGANIZAÇÃO E REPRESENTAÇÃO
DO CONHECIMENTO

SUMÁRIO

PREFÁCIO	6
GT 2 – ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO.....	334
Modalidade da apresentação: Comunicação oral	334
ESTUDO DA RELAÇÃO ENTRE ORGANIZAÇÃO E VISUALIZAÇÃO DA INFORMAÇÃO A PARTIR DE SISTEMAS DE RECUPERAÇÃO DE INFORMAÇÃO	334
<i>Jéssica Monique de Lira Vieira</i> <i>Fábio Assis Pinho</i>	
ANÁLISE DE ASSUNTO EM BIBLIOTECAS UNIVERSITÁRIAS: A OBSERVAÇÃO DE ESTRUTURAS TEXTUAIS EM LIVROS DAS ÁREAS DE HUMANAS, EXATAS E BIOLÓGICAS	353
<i>Daniela Majorie dos Reis</i> <i>Mariângela Spotti Lopes Fujita</i>	
MODELAGEM E REPRESENTAÇÃO SEMÂNTICA DE DADOS GOVERNAMENTAIS ABERTOS DA PREVIDÊNCIA SOCIAL BRASILEIRA	372
<i>Durval Vieira Pereira</i> <i>Carlos Henrique Marcondes</i>	
TEORIA DA TERMINOLOGIA E TEORIA DA ONTOLOGIA: UM COMPARATIVO BASEADO NA CRIAÇÃO DE DEFINIÇÕES PARA TERMOS	394
<i>Elisângela Cristina Aganette</i> <i>Mauricio Barcellos Almeida</i>	
A SEMIÓTICA COMO SUBSÍDIO PARA A REPRESENTAÇÃO DO CONHECIMENTO: UMA ANÁLISE CONCEITUAL SOBRE O TEMA	Erro!
Indicador não definido. <i>Isadora Victorino Evangelista</i> <i>José Augusto Chaves Guimarães</i> <i>Carlos Cândido Almeida</i>	
PRINCÍPIOS METODOLÓGICOS PARA DESENVOLVIMENTO DE ONTOLOGIAS: ANÁLISE DAS PRÁTICAS CORRENTES E PROPOSIÇÃO DE MELHORIAS.....	430
<i>Fabício Martins Mendonça</i> <i>Maurício Barcellos Almeida</i>	
A ANÁLISE TIPOLÓGICA COMO SUBSÍDIO PARA CONSTRUÇÃO DE VOCABULÁRIO CONTROLADO: O CASO DO JUÍZO DE ÓRFÃOS DO RECIFE (1824-1889)	448
<i>Mônica Maria Pádua Souto da Cunha</i> <i>Fábio Mascarenhas Silva</i>	
ENTRE O USUÁRIO E A INFORMAÇÃO CORPORATIVA: A BUSCA POR FACETAS.....	468
<i>Leonardo Lacerda Alves</i>	

ÍCONE, ÍNDICE E SÍMBOLO, FUNDAMENTOS PARA LER E ORGANIZAR A INFORMAÇÃO EM IMAGENS	488
<i>Paulo Roberto Gomes Pato</i>	
PADRÕES LINGUÍSTICOS PARA EXTRAÇÃO AUTOMÁTICA DO OBJETIVO .	509
<i>Leonardo Cruz Costa</i>	
<i>Luciana Reis Malheiros</i>	
PERSONALIDADE E MATÉRIA NA TEORIA DA CLASSIFICAÇÃO FACETADA: A QUESTÃO DO CONTEXTO, PRESSUPOSTOS TEÓRICOS E METODOLÓGICOS	528
<i>Linair Maria Campos</i>	
<i>Maria Luiza de Almeida Campos</i>	
MAPEAMENTO TEMÁTICO: O GT2 DO ENANCIB.....	549
<i>Rosana Portugal Tavares de Moraes</i>	
<i>Maria Luiza de Almeida Campos</i>	
DA MÚSICA À ORGANIZAÇÃO DO ACERVO: UMA ANÁLISE DA REPRESENTAÇÃO DA INFORMAÇÃO DA BIBLIOTECA JOSÉ ALBERTO KAPLAN	569
<i>Ana Claudia Medeiros de Sousa</i>	
<i>Maria Elizabeth Baltar Carneiro de Albuquerque</i>	
ORGANIZAÇÃO DOS ESTUDOS INTERDISCIPLINAR DE SUBSTÂNCIAS PSICOATIVAS	584
<i>Miriam Gontijo Moraes</i>	
IDENTIFICAÇÃO DE TIPOLOGIA DOCUMENTAL COMO PARÂMETRO PARA CLASSIFICAÇÃO	601
<i>Silvia Lhamas de Mello</i>	
<i>Ana Célia Rodrigues</i>	
ANÁLISE E DESCRIÇÃO DOS METADADOS PRESENTES NOS ELEMENTOS DA MODELIZAÇÃO DE UM DOMÍNIO ORGANIZACIONAL.....	621
<i>Célia da Consolação Dias</i>	
<i>Lídia Alvarenga</i>	
A ORGANIZAÇÃO DO CONHECIMENTO ARQUIVÍSTICO: O PRINCÍPIO DA PROVENIÊNCIA E O SISTEMISMO	639
<i>Josemar Henrique de Melo</i>	
<i>Américo Augusto Nogueira Vieira</i>	
<i>Guilherme Ataíde Dias</i>	
OS REGISTROS DE AUTORIDADE EM SISTEMAS DE INFORMAÇÃO: UMA PERSPECTIVA BIBLIOTECÔNOMICA E ARQUIVÍSTICA	660
<i>Dunia Llanes Padrón</i>	
<i>Mariângela Spotti Lopes Fujita</i>	
<i>Flávia Maria Bastos</i>	
PRAGMÁTICA NA ORGANIZAÇÃO DO CONHECIMENTO	681
<i>Leila Cristina Weiss</i>	
<i>Marisa Bräscher</i>	

DIMENSÕES SUBJETIVAS NO TRATAMENTO DA INFORMAÇÃO: ANALISANDO O PROCESSO DE INDEXAÇÃO E IMPLICAÇÕES SUBJACENTES NA RECUPERAÇÃO DA INFORMAÇÃO.....	702
<i>Eliane Pawlowski de Oliveira Araujo</i>	
<i>Claudio Paixão Anastácio de Paula</i>	
ORGANIZAÇÃO DA INFORMAÇÃO JURÍDICA: O ARQUIVO DA JUSTIÇA FEDERAL DA PARAÍBA (JFPB), USOS E USUÁRIOS.....	721
<i>Wendia Oliveira de Andrade</i>	
<i>Dulce Amelia de Britto Neves</i>	
IDENTIFICAÇÃO DE PRINCÍPIOS UTILITARISTAS DE JEREMY BENTHAM NA CONSTRUÇÃO DE SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO	740
<i>Marcos Luiz Cavalcanti de Miranda</i>	
<i>Icléia Thiesen</i>	
MODELO DE COLABORAÇÃO PARA INDEXAÇÃO DE RECURSOS WEB	759
<i>Marcel Ferrante Silva</i>	
<i>Gercina Angela Borém de Oliveira Lima</i>	
ABORDAGEM TEÓRICO-METODOLÓGICA NA ORGANIZAÇÃO DE IMAGENS EM PATOLOGIA.....	778
<i>Joice Cleide Cardoso Ennes de Souza</i>	
<i>Elan Cardozo Paes de Almeida</i>	
A REPRESENTAÇÃO DA INFORMAÇÃO: CLASSIFICAÇÃO E INDEXAÇÃO AUTOMÁTICA DE DOCUMENTOS DE ARQUIVO.....	798
<i>Renato Tarciso Barbosa de Sousa</i>	
A NOÇÃO DE DOCUMENTAÇÃO EM MUSEUS NAS NORMAS SPECTRUM E CIDOC/ICOM	812
<i>Juliana Monteiro</i>	
<i>Marilda Ginez Lopes de Lara</i>	
IDENTIFICAÇÃO ARQUIVÍSTICA COMO REQUISITO DA GESTÃO DE DOCUMENTOS: CONTRIBUIÇÃO PARA APLICAÇÃO DA LEI DE ACESSO À INFORMAÇÃO NO BRASIL.....	830
<i>Ana Célia Rodrigues</i>	
ORGANIZAÇÃO E REPRESENTAÇÃO DA INFORMAÇÃO NO MUSEU VIRTUAL DE INSTRUMENTOS MUSICAIS – MVIM.....	848
<i>Adriana Olinto Ballesté</i>	
<i>Alea de Almeida</i>	
A SEMIÓTICA DA CULTURA NAS ABORDAGENS SOCIOCULTURAIS DA ORGANIZAÇÃO DO CONHECIMENTO: UMA ANÁLISE TEÓRICO-CONCEITUAL	Erro! Indicador não definido.
<i>Mona Cleide Quirino da Silva Farias</i>	
<i>Carlos Cândido de Almeida</i>	
TERMINOLOGIA E DOCUMENTAÇÃO: ESTUDO TERMINOGRÁFICO DA PERFORMANCE MUSICAL.....	893
<i>Soraia de Andrade Lara Carvalho</i>	
<i>Gercina Ângela B. de Oliveira Lima</i>	

ANÁLISE DE ONTOLOGIAS PARA REPRESENTAÇÃO DE DOCUMENTOS MULTIMÍDIA NA WEB	913
<i>Daniela Lucas Silva</i>	
<i>Renato Rocha Souza</i>	
ESTUDO COMPARATIVO DE ESTRUTURAS DOCUMENTAIS EM INSTITUIÇÕES ANÁLOGAS: UMA ANÁLISE TIPOLOGICA SOBRE INVENTÁRIOS POST-MORTEM.....	934
<i>Sonia Maria Troitiño Rodriguez</i>	
MINERAÇÃO DE TEXTOS PARA A EXTRAÇÃO AUTOMÁTICA DE SINTAGMAS NOMINAIS: TENDÊNCIAS E ESTRATÉGIAS PARA RECUPERAÇÃO DA INFORMAÇÃO JURÍDICA	954
<i>Fanny do Couto Ribeiro de Lima</i>	
CLASSIFICAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO EM CT&I NO CONTEXTO DA DIVULGAÇÃO CIENTÍFICA EM FONTES DE INFORMAÇÃO DIGITAIS	974
<i>Arlete Sandra Mariano Alves Baubier</i>	
<i>Rosali Fernandez de Souza</i>	
<i>Lena Vania Ribeiro Pinheiro</i>	
INSTITUCIONALIZAÇÃO CIENTÍFICA DA ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO NA CIENCIA DA INFORMAÇÃO NO BRASIL	993
<i>Gracy Kelli Martins</i>	
<i>João Batista Ernesto de Moraes</i>	
PROPOSTA DE MODELAGEM CONCEITUAL A PARTIR DA IMPLEMENTAÇÃO DA TEORIA DA CLASSIFICAÇÃO FACETADA DE RANGANATHAN EM BANCO DE DADOS	1011
<i>Márcio Bezerra Da Silva</i>	
<i>Zeny Duarte de Miranda</i>	
Modalidade da apresentação: Pôster.....	1028
APLICAÇÃO DE MODELO DE REENGENHARIA EM TESAURO	1028
<i>Benildes Coura Moreira dos Santos Maculan</i>	
<i>Gercina Ângela Borem de Oliveira Lima</i>	
CIÊNCIA DA INFORMAÇÃO, POSITIVISMO E MEMÓRIA: RELAÇÕES COM A REDOCUMENTARIZAÇÃO DO DOCUMENTO DE VALOR HISTÓRICO	1037
<i>Hercules Pimenta dos Santos</i>	
ANÁLISE DE DOMÍNIO HJØRLANDIANA SOB A LUZ DA FILOSOFIA DE DELEUZE	1044
<i>Igor Soares Amorim</i>	
<i>Lígia Maria Arruda Café</i>	
METODOLOGIAS, INSTRUMENTOS E INTERFACES DE ORGANIZAÇÃO DA IMAGEM NA WEB NO CONTEXTO DO SISTEMA “MEMÓRIA VIRTUAL”	1051
<i>Letícia Reis da Silveira</i>	
<i>Luciana de Souza Gracioso</i>	
ORGANIZAÇÃO DA INFORMAÇÃO E DO CONHECIMENTO DE DOCUMENTOS ARTÍSTICOS: INTEGRANDO MUSEU, BIBLIOTECA E ARQUIVO	1058
<i>Giovana Deliberali Maimone</i>	

VARIÁVEIS DA ANÁLISE DOCUMENTÁRIA DE CHARGE: UMA PROPOSTA INICIAL.....	1066
<i>Thulio Pereira Dias Gomes</i>	
<i>Rosali Fernandez de Souza</i>	
NANOPUBLICAÇÕES E INDEXAÇÃO: ASPECTOS METODOLÓGICOS.....	1073
<i>Lorena Tavares de Paula</i>	
<i>Maria Aparecida Moura</i>	
CATEGORIZAÇÃO DE PATENTES COM USO DE INTELIGÊNCIA COMPUTACIONAL.....	1081
<i>Magali Rezende Gouvêa Meireles</i>	
<i>Beatriz Valadares Cendón</i>	
A FOLKSONOMIA E A REPRESENTAÇÃO COLABORATIVA DA INFORMAÇÃO EM AMBIENTES DIGITAIS	1089
<i>Raimunda Fernanda dos Santos</i>	
<i>Renato Fernandes Corrêa</i>	
TESAURO DO REGIME MILITAR BRASILEIRO.....	1097
<i>Renato Rocha Souza</i>	
<i>Suemi Higuchi</i>	
<i>Alexandre Morelli</i>	
<i>Angela Moreira</i>	
<i>Bernardo de Paola Bortolotti Faria</i>	
<i>Bianca de Magalhães Silveira</i>	
<i>Celso Castro</i>	
<i>Daniele Chaves Amado</i>	
<i>Juliana Gagliardi</i>	
<i>Letícia Carvalho Ferreira</i>	
<i>Marco Aurélio Vannuchi Leme de Mattos</i>	
<i>Martina Spohr</i>	
<i>Renan Marinho de Castro</i>	
<i>Vanessa Matheus Cavalcante</i>	
O TRATAMENTO DESCRITIVO E TEMÁTICO DE ACERVOS FOTOGRÁFICOS NO PARANÁ	1107
<i>Ana Cristina de Albuquerque</i>	
<i>Ana Carolina Simionato</i>	
EXAUSTIVIDADE, PRECISÃO E CONSISTÊNCIA EM INDEXAÇÃO: SISTEMATIZAÇÃO CONCEITUAL	1114
<i>Luciana Beatriz Piovezan</i>	
<i>Mariângela Spotti Lopes Fujita</i>	
A ARQUITETURA DE METADADOS DE PROCESSOS JUDICIAIS CÍVEIS	1122
<i>Eduardo Watanabe</i>	
A CIÊNCIA DIPLOMÁTICA PARA UMA ANÁLISE DE DOCUMENTOS FOTOGRÁFICOS: UMA REFLEXÃO PAUTADA NA PUBLICIDADE	1129
<i>Aline Ribeiro de Farias</i>	
<i>Telma Campanha de Carvalho Madio</i>	
O VOCABULÁRIO CONTROLADO COMO FERRAMENTA DO PROCESSO DE ORGANIZAÇÃO E RECUPERAÇÃO DA INFORMAÇÃO.....	1136

Luciana Davanzo
Walter Moreira

TESAURO AFRO-BRASILEIRO: USO ESTRATÉGICO PARA ORGANIZAÇÃO E
RECUPERAÇÃO DE INFORMAÇÃO 1143

Graziela dos Santos Lima
Marcia Silveira Kroeff
Divino Ignacio Ribeiro Junior

ANÁLISE DOCUMENTÁRIA DE IMAGENS FOTOGRÁFICAS DO ARQUIVO
AFONSO PEREIRA 1151

Eveline Filgueiras Gonçalves
Maria Elizabeth Baltar Carneiro de Albuquerque

ORGANIZAÇÃO DA INFORMAÇÃO EM CONTEXTO DINÂMICO E SENSÍVEL:
TESAURO DE INTELIGÊNCIA POLICIAL 1158

Maria Aparecida Moura
Rodrigo Benjamin Gesteira

PREFÁCIO

A Ciência da Informação é um campo científico de natureza interdisciplinar devotado à busca por soluções para a efetiva comunicação da informação, bem como de seus registros, [contexto social não é entre pessoas?] no contexto social, institucional ou individual de uso e a partir de necessidades específicas. A evolução da Ciência da Informação está inexoravelmente ligada à tecnologia da informação, uma vez que o imperativo tecnológico tem gerado transformações que culminaram em uma sociedade pós-industrial, a sociedade da informação. Nesse contexto, a Ciência da Informação desempenha importante papel na evolução da sociedade da informação por suas fortes dimensões social e humana, as quais vão além das fronteiras da tecnologia.

O tema do ENANCIB 2014 – *Além das nuvens: expandindo as fronteiras da Ciência da Informação* – remete ao cenário atual caracterizado pelo contínuo desenvolvimento das tecnologias da informação e comunicação, assim como pela evolução constante do ambiente *Web*, os quais têm proporcionado novas formas de acessar, recuperar, armazenar e gerir a informação. Telefonia móvel, nuvens, *big data*, *linked data*, dentre outras formas de interagir com a informação têm exigido novas abordagens para os estudos em Ciência da Informação. O ENANCIB 2014 oferece a oportunidade para refletir sobre essas mudanças, as quais impactam na interação humana com a informação, bem como sobre suas implicações para o futuro da Ciência da Informação.

Promovido pela Associação Nacional de Pesquisa em Ciência da Informação (ANCIB), o ENANCIB, em sua décima quinta edição, foi organizado pelo Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Minas Gerais (PPGCI-ECI/UFMG) e realizado na Escola de Ciência da Informação da Universidade Federal de Minas Gerais (ECI/UFMG), em Belo Horizonte, Minas Gerais, no período de 27 a 31 de outubro de 2014. O evento foi financiado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), pela UFMG e outras organizações apoiadoras.

Pesquisadores em Ciência da Informação foram convidados a submeter pesquisas teóricas e empíricas, de acordo com a orientação temática dos onze Grupos de Pesquisa (GTs) da ANCIB. A chamada de trabalhos foi aberta para duas categorias de submissões. A primeira categoria é a comunicação oral (máximo de 20 páginas), que consiste de artigo escrito em português, descrevendo trabalho original com demonstração efetiva de resultados. As

comunicações orais aprovadas foram convidadas para apresentação no evento. A segunda categoria é o pôster (máximo de 7 páginas), que consiste de artigos curtos escritos em português, descrevendo pesquisa em desenvolvimento. Os pôsters aceitos foram convidados para exposição nas dependências em que ocorreu o evento.

O ENANCIB 2014 recebeu mais de 600 trabalhos, dos quais mais de 300 foram aceitos para publicação nos Anais, sendo cerca de 240 para apresentação oral e 80 para exibição em pôsters. Este volume é então constituído por 74% de comunicações orais e 26% de pôsteres, selecionados pelo comitê de programa dos GTs, os quais são compostos por pareceristas especializados, definidos no âmbito de cada GT.

Agradecemos à Comissão Organizadora e à ANCIB pelo seu comprometimento com o sucesso do evento, aos autores por suas submissões e à Comissão Científica pelo intenso trabalho. Agradecemos ainda aos alunos, funcionários e colaboradores que contribuíram para a efetivação do evento.

Belo Horizonte, outubro de 2014

Isa M. Freire
Lilian M. A. R. Alvares
Renata M. A. Baracho
Mauricio B. Almeida
Beatriz V. Cendon
Benildes C. M. S. Maculan

GT 2 – ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO

Modalidade da apresentação: Comunicação oral

ESTUDO DA RELAÇÃO ENTRE ORGANIZAÇÃO E VISUALIZAÇÃO DA INFORMAÇÃO A PARTIR DE SISTEMAS DE RECUPERAÇÃO DE INFORMAÇÃO

STUDY OF THE RELATIONSHIP BETWEEN ORGANIZATION AND VISUALIZATION INFORMATION FROM INFORMATION RETRIEVAL SYSTEMS

Jéssica Monique de Lira Vieira
Fábio Assis Pinho

Resumo: Buscando propor estratégias que potencializem o entendimento da memória armazenada em Sistemas de Recuperação de Informação, considera-se como alternativa a adoção de procedimentos de Organização da Informação e Visualização da Informação. Nesse contexto, esta pesquisa tem como objetivo analisar e identificar a relação entre os aspectos da Organização e da Visualização da Informação do ponto de vista da contribuição das duas áreas para os Sistemas de Recuperação de Informação, razão pela qual a pesquisa é caracterizada como exploratória e bibliográfica. O percurso metodológico utilizado para realização da pesquisa consiste em levantamento bibliográfico, para obtenção do embasamento científico para melhor compreensão e análise de cada área; levantamento de sistemas que fazem uso de interface visual em sua estrutura, bem como da análise dos elementos de organização da informação necessários para a representação de conteúdos a serem transmitidos em estruturas visuais. A partir disso, evidenciou-se a interligação entre as áreas abordadas, quando aplicadas em sistemas que lidam com a recuperação da informação. Desse modo, concluiu-se que a Organização e a Visualização da Informação podem alcançar promissores resultados se atuarem em parceria para otimização na recuperação da memória armazenada em Sistemas de Recuperação de Informação.

Palavras-chave: Organização da Informação. Visualização da Informação. Sistemas de Recuperação de Informação.

Abstract: Trying to propose strategies that enhance the understanding of memory stored in Information Retrieval Systems, this thesis considers, as an alternative, the adoption of procedures of Information Organization and Information Visualization. Thus, this research aims to analyze and identify the relationship between the resources from the Information Organization and Visualization from the point of view of the contribution of the two areas for Information Retrieval Systems, for that reason the research is characterized as exploratory and bibliographical. The methodological procedure used to conduct the research consists of bibliographic survey, for obtaining scientific basis for better understanding and analysis of each area; a survey of systems that make use of visual interface in its structure, as well as the analysis of the information organization criteria necessary to representation of content to be transmitted in visual structures. From that, it was evident the interconnection between the areas addressed when applied in systems that deal with information retrieval. Thus, it is concluded that Information Organization and Visualization can achieve promising results if they work in partnership to optimize the recovery of memory stored in Information Retrieval Systems.

Keywords: Information organization. Information Visualization. Information Retrieval Systems.

1 INTRODUÇÃO

A crescente produção de informações, juntamente com os avanços tecnológicos, desempenha um papel fundamental na consolidação da sociedade atual, enquanto “sociedade da informação”. Sociedade em que a informação é vista, como afirma Le Coadic (2004), como uma necessidade fundamental, ou seja, como uma necessidade vital para o desenvolvimento de, praticamente, todas as atividades humanas. É nesse cenário, no qual a informação, dotada de tamanho valor, enfrenta um crescimento na sua produção e disponibilização, que se faz necessária à busca de métodos adequados para, significativamente, gerenciar a quantidade de informação disponível de forma a possibilitar seu acesso e uso.

Diante desse contexto e da necessidade de proporcionar o acesso para posterior uso da informação, torna-se salutar propor soluções para os problemas relacionados à informação e à sua recuperação. Nesse âmbito, uma das áreas preocupadas e envolvidas com essa questão é a Ciência da Informação (CI) que, juntamente com outras áreas, é uma participante ativa e deliberada na evolução da sociedade da informação (SARACEVIC, 1995); e, como aponta Dias (2001), tem como questão básica o acesso à informação, ou a facilitação desse acesso, o que constitui uma das razões gerais de sua existência.

Assim, a CI, nas palavras de Alves *et al.* (2007, p. 29, grifo nosso), consiste em uma área que:

busca investigar, entre outros assuntos, a geração, a **organização**, o processamento, a **recuperação** e o uso da informação, em diversos ambientes. É, por natureza, uma área interdisciplinar [...] Além disso, está intimamente ligada com as Tecnologias de Informação e Comunicação vigentes.

Nessa perspectiva, entre as áreas estudadas pela CI, está a Recuperação da Informação (RI), que se preocupa em disponibilizar a informação para satisfação de uma necessidade do usuário. No âmbito da RI, torna-se oportuno evidenciar o papel dos sistemas que atuam no intuito de tornar possível, a partir de um universo de documentos disponíveis, a recuperação daqueles que venham a ser úteis para quem deles necessita, os chamados Sistemas de Recuperação de Informação (SRIs). Tais sistemas, como destaca Araújo (1995), constituem a “memória humana registrada”.

No entanto, para que os SRIs atuem de maneira satisfatória, é fundamental que haja uma atenção maior no que tange à organização das informações armazenadas por esses sistemas, pois, como destaca Brascher e Carlan (2010), em um sistema de informação, a

qualidade obtida na recuperação da informação depende, substancialmente, dos procedimentos e instrumentos utilizados para organização da informação.

Nesse contexto, tem-se a Organização da Informação (OI), que exerce uma atividade nuclear enquanto mediadora, sendo responsável por estabelecer uma ponte entre a representação e o uso da informação, caracterizando, assim, seu papel, quando se lida com objetos informacionais. A OI tem como função, resumidamente, organizar registros de informação para arranjá-los sistematicamente, de forma a possibilitar sua recuperação e posterior uso.

Para desenvolvimento de suas práticas, a OI fundamenta-se em um aspecto essencial, o tratamento da informação. O tratamento da informação, de acordo com Pinto e Molina (1993 apud DAL'EVEDOVE, 2011), consiste no processo de gerar representações documentais, as quais, inseridas no contexto de sistemas de informação, atuam como instrumentos de busca, cujo propósito corresponde à facilitação da consulta e disseminação do documento primário.

Considerando essa realidade, ressalta-se o importante valor que a OI representa para os sistemas de informação, tendo em vista que a adequada organização tem o poder de interferir diretamente na recuperação de informações precisas e que, em contrapartida, a informação sem uma organização prévia pode permanecer oculta e sem uso. Reafirmando essa ideia, Svenonius (2000) enfatiza, ainda, que a eficácia de um sistema que lida com a recuperação da informação está diretamente relacionada à inteligência adotada para a sua organização.

Outro ponto a ser tratado quando se lida com a otimização na recuperação e uso da memória armazenada nos sistemas de informação refere-se à Visualização da Informação (VI). A VI pode ser vista como uma área que tem como finalidade potencializar a apropriação de informação, por meio de recursos gráficos conhecidos como estruturas de visualização da informação (DIAS; CARVALHO, 2007).

A VI apóia-se na representação visual como maneira de auxiliar a compreensão de informações armazenadas e que foram, previamente, organizadas, objetivando facilitar a localização e recuperação eficaz de itens que atendam a uma necessidade. Para isso, a VI faz uso de interfaces gráficas que possibilitam a interação do usuário com os sistemas, por meio de ambiente de pesquisa agradável, com a finalidade de maximizar o uso da informação.

No entanto, para que a VI possa exercer sua função de forma satisfatória, é preciso que os procedimentos de OI ofereçam um suporte que torne possível a adoção de representações visuais pelos sistemas (ZHANG, 2008), pois a OI se configura como área fundamental na

organização e tratamento da informação, e que afeta diretamente sua recuperação; por outro lado, a VI se configura como uma poderosa aliada, capaz de potencializar o acesso, localização e recuperação da informação organizada. Assim, tem-se a OI como um processo interno aos sistemas de recuperação de informação, porém, muitas vezes, não transparente; e a VI que, atuando mediante representação visual, visa à facilitação na apropriação das informações previamente estruturadas e organizadas.

É nesse cenário que se estabelece o objetivo deste trabalho, analisar e identificar a relação entre os aspectos da OI e da VI do ponto de vista da contribuição das duas áreas para os SRIs, de maneira a potencializar o entendimento da memória por eles armazenada.

2 ORGANIZAÇÃO DA INFORMAÇÃO

O verbo “organizar”, de acordo com o Dicionário Houaiss, significa “dar determinada ordem a” (série de itens); “dispor de forma ordenada”; “dar forma regular às partes de” (um todo); “arrumar, ordenar”. Assim também se comporta o ser humano, organizando naturalmente os muitos tipos de coisas em sua vida: os livros em estantes, as roupas em armários, utensílios de cozinha em gavetas, documentos específicos em pastas, etc. Essa organização, mesmo que inconscientemente, visa à facilitação de uma futura localização.

No contexto dos sistemas de informação não é diferente. Os sistemas organizam um acervo, cuja função é armazenar objetos informacionais, compreendendo melhor de que tratam, dispendo de forma ordenada, de maneira que as informações neles registradas possam ser recuperadas com facilidade. Essa prática de armazenar e organizar o conteúdo informacional produzido, visando alcançar usuários futuros, não é recente, como salienta McGarry (1999, p. 112): “[...] o anseio de armazenar os produtos da memória coletiva tem estado conosco desde o início da era da escrita”.

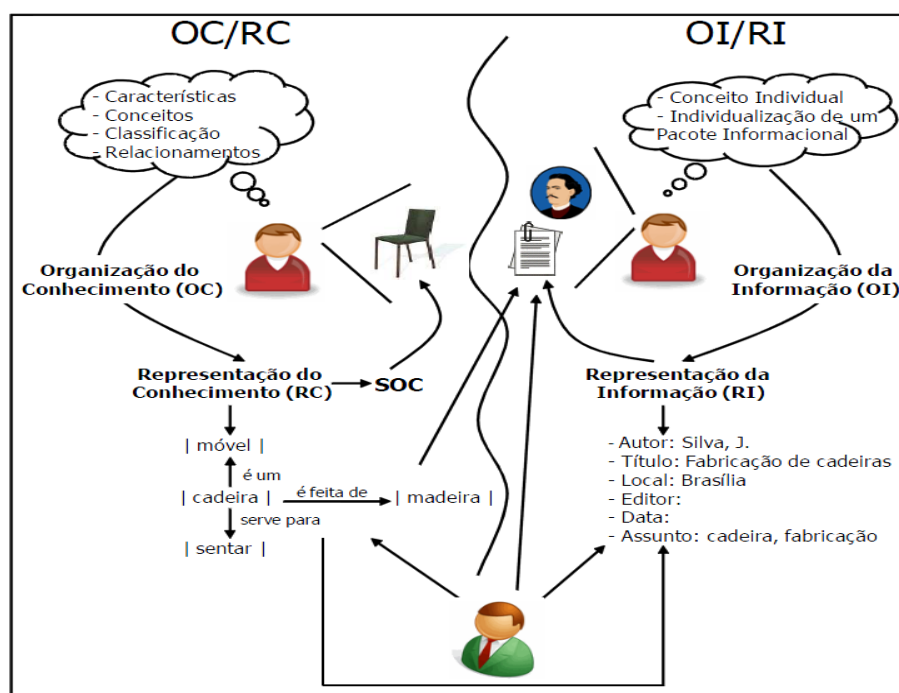
Antes de prosseguir com a abordagem acerca do conceito de Organização da Informação, seu significado e práticas, é importante diferenciá-lo do conceito de Organização do Conhecimento. Assim, Brascher e Café (2008), que discorrem em seu trabalho sobre a divergência entre os conceitos, abordam a existência de uma discordância entre a delimitação de Organização da Informação (OI) e Organização do Conhecimento (OC) que, por vezes, são utilizados com falta de clareza em seus sentidos.

Para as autoras, a OI se aplica às ocorrências individuais de objetos informacionais, compreendendo a organização desses objetos para arranja-los sistematicamente em coleções. A OC, por sua vez, se aplica às unidades do pensamento (conceitos), e visa a construção de

modelos de mundo que se constituem em abstrações da realidade, que permitem descrever e fornecer explicações sobre os fenômenos observados.

De forma resumida, a OI tem como objeto a informação registrada passível de ser armazenada e organizada, possibilitando sua recuperação quando necessário. Já a OC se encontra no mundo dos conceitos que visam à construção de representações do conhecimento. A FIG. 2, a seguir, ilustra essa proposta em relação a diferenciação entre os dois conceitos.

FIGURA 1 – OC/RC e OI/RI



Fonte: Brascher e Café (2008, p.7)

Com base no exposto acerca da OI e da OC, optou-se aqui pela utilização do termo Organização da Informação, já que sua proposta se adequa ao objetivo que se pretende alcançar com esta pesquisa, ou seja, analisar os aspectos da OI e sua influência na prática da organização que resulta na facilitação do acesso e recuperação de documentos informacionais em SRIs.

Dessa forma, considera-se a proposta conceitual de Brascher e Café (2008), em que a OI compreende a organização de um conjunto de objetos informacionais para arranjá-los sistematicamente em coleções, como a organização da informação em bibliotecas, museus, arquivos, tanto tradicionais quanto eletrônicos. Considera-se, também, a contribuição de Alvarenga (2003), que acrescenta ainda que os processos de organização da informação destinam-se, prioritariamente, à recuperação eficaz de informações por parte dos usuários.

Assim, para melhor compreender como se desenvolvem as práticas da OI, se faz necessário uma análise mais aprofundada a respeito de um aspecto essencial que fundamenta suas atividades - o tratamento da informação.

2.1 Tratamento da informação

Contextualizando o tratamento da informação em sistemas de informação, Dias (2001, p.3) o define como “a função de descrever os documentos, tanto do ponto de vista físico (características físicas dos documentos) quanto do ponto de vista temático (ou de descrição do conteúdo)”. Sendo assim, o tratamento da informação busca representar os documentos, por meio da síntese deles, para que sua organização e manipulação dentro do sistema de recuperação possam ser facilitadas e para que o usuário possa avaliar, de maneira mais simples, a relevância que os documentos apresentam para sua necessidade de informação.

Nessa linha de pensamento, Foskett (1973 apud GUIMARÃES, 2009) reconhece que a OI é uma atividade de natureza eminentemente mediadora, em que dois universos se descortinam: o primeiro, ligado ao acesso físico aos documentos; e o segundo, de natureza mais complexa, voltado para o acesso ao conteúdo informacional. Desse modo, a organização da informação é composta por dois grandes processos que estão relacionados entre si: o tratamento descritivo e o tratamento temático.

Após análise da OI e dos processos que fundamentam as suas atividades, se faz relevante abordar a VI, enquanto área que contribui, por meio da construção de interfaces gráficas, para representação visual de conteúdo informacional e atua no intuito de facilitar o acesso às informações que foram previamente organizadas.

3 VISUALIZAÇÃO DA INFORMAÇÃO

A VI é definida como uma área da ciência que tem por objetivo o estudo das principais formas de representações gráficas para apresentação de informações, a fim de contribuir para o melhor entendimento delas, bem como ajudar a percepção do usuário na dedução de novos conhecimentos baseado no que está sendo apresentado (FREITAS *et al.*, 2001).

Outra interessante definição é trazida pelo grupo de pesquisa em VI do instituto de pesquisa de *software* da Universidade da Califórnia, Irvine¹, segundo o qual a VI está centrada no desenvolvimento e análise de métodos para a apresentação de informações abstratas de forma visual. Tendo em vista que a representação visual de informações permite que as pessoas se tornem mais facilmente consciente de fatos essenciais, desenvolvendo uma

¹ Disponível em: <<http://www.isr.uci.edu/research.html>>.

compreensão mais profunda sobre o que está sendo apresentado e permitindo, por meio de uma visualização interativa, manipular e explorar informações.

Resumidamente, a VI está fundamentada no fato de que a informação apresentada de maneira visual possibilita uma compreensão mais clara daquilo que está sendo transmitido, já que, como acrescenta Dias e Carvalho (2007, p.1):

na maioria dos casos o oferecimento de imagens, figuras, estruturas gráficas e quaisquer outros recursos gráficos, com a finalidade de apresentar uma informação, produz a compreensão da mensagem transmitida, pois esta se torna mais natural e exige menos esforço cognitivo.

É importante entender, também, que a VI, tendo como objetivo final a inclusão informacional dos usuários, conta, de maneira geral, com o auxílio da tecnologia e dos computadores para geração de representações gráficas, se apoiando no uso da capacidade visual humana que, como afirma Nascimento e Ferreira (2005), é o sentido que possui maior capacidade de captação de informações por unidade de tempo.

Como visto até o momento, a VI, enquanto atividade temática, está voltada para as questões que envolvem a otimização na representação e, conseqüentemente, no acesso, localização e recuperação de informações.

Buscando compreender de forma mais clara a influência que a visualização exerce na recuperação de informação, se faz relevante destacar a visão de Zhang (2008) ao elucidar que o dinamismo, a diversidade, a heterogeneidade e a complexidade de informações dispostas na internet e em outros sistemas de recuperação de informação têm representado um desafio para as técnicas e teorias tradicionais de recuperação de informação. Esses desafios levam a necessidade de adoção de uma maneira mais interativa, intuitiva e eficaz de recuperação de informação nos sistemas. Nessa perspectiva, encontra-se a VI que, por meio das interfaces gráficas, se configura como uma poderosa aliada em SRIs.

Assim, quando tais sistemas possuem suas informações devidamente organizadas, a utilização de interfaces gráficas pode vir a ser um potencializador na representação e, conseqüentemente, na recuperação, minimizando algumas das dificuldades enfrentadas pelos sistemas, já que o uso dessas interfaces facilita a interação com o usuário e permite, como afirma Mazza (2009), entender sistemas complexos, tomar decisões e descobrir informações importantes que, de outra maneira, poderiam permanecer ocultas.

4 METODOLOGIA

Nesta seção apresenta-se a metodologia que caracteriza a pesquisa, bem como o percurso metodológico utilizado para sua realização.

4.1 Tipo de Pesquisa

Esta pesquisa se caracteriza como exploratória, uma vez que, para o alcance dos objetivos propostos, buscou-se levantar e explorar referencial teórico acerca da organização e representação visual de informações como alternativa para potencializar o acesso e a recuperação do conteúdo informacional armazenado em SRIs. Assim, por meio da pesquisa exploratória, que, de acordo com Gil (2009, p. 27) “tem como principal finalidade desenvolver, esclarecer [...] conceitos e ideias, tendo em vista a formulação de problemas mais precisos”, buscou-se proporcionar uma maior familiaridade com as áreas abordadas.

Com o intuito de possibilitar a formação de um referencial teórico acerca das temáticas aqui estudadas, a pesquisa se caracteriza, também, como bibliográfica, pois, a partir da pesquisa bibliográfica, obtém-se o embasamento científico para melhor compreensão e análise de cada área.

4.2 Levantamento Bibliográfico

Inicialmente, foi feito um levantamento bibliográfico baseado em literatura voltada para as temáticas aqui abordadas, com o intuito de fundamentar a pesquisa, e, para isso, foram analisadas referências teóricas extraídas de artigos, teses, dissertações e livros, presentes na literatura nacional e internacional, que versem sobre Organização da Informação, Visualização da Informação e Sistemas de Recuperação de Informação. As buscas por artigos foram realizadas, também, nas Bases de Dados: BRAPCI e SCIELO; e no portal de periódicos da CAPES.

4.3 Descrição de técnicas e modelos de interface de VI

Em seguida, buscou-se abordar as técnicas e modelos de interface de VI, sendo analisados aqueles que melhor se adequam aos SRIs, ou seja, que tenham aplicações práticas na construção de interfaces para sistemas que lidam com a recuperação da informação, possibilitando a representação e visualização da estrutura da coleção de documentos.

Nessa perspectiva, quanto às técnicas, foram analisadas: - Desenho de Grafos; - Desenho de Árvores; - *Browser* Hiperbólico.

Quanto aos modelos de VI foram analisados os seguintes: - Mapa Conceitual; - Cartografia Temática; - Nuvem de *Tags*.

4.4 Análise de sites de SRIs

Por último, foi feito um levantamento de SRIs na Web que adotam visualizações na interface com o usuário. Cada um dos sistemas foi escolhido por fazer uso de uma técnica ou

modelo de interface visual anteriormente abordado, e por ter como finalidade a representação e recuperação de documentos como registros de informação.

Nesse contexto, apresenta-se uma análise de como deveriam estar organizadas e estruturadas as informações e quais instrumentos seriam necessários para tornar possível a aplicação eficaz de cada técnica e modelo de VI nos sistemas. Para a análise, foram abordados os seguintes *sites* de SRIs e suas interfaces visuais:

- **AquaBrowser da biblioteca da Universidade de Edimburgo** - o catálogo on line da biblioteca faz uso de interface baseada na técnica desenho de grafos;
- **Mecanismo de busca YIPPY** - faz uso do desenho de árvores de diretórios para organizar os documentos retornados para uma consulta realizada em sua base de documentos;
- **Infoteca da Embrapa** – faz uso da técnica Browser Hiperbólico para permitir acesso a informações sobre tecnologias produzidas pela Embrapa;
- **Website do IHMC Camp Tools** - utiliza o mapa conceitual como modelo de interface para descrever o programa Cmap Tools;
- **WEBSOM** - método para construção de mapas visuais – utiliza a cartografia temática como um método para gerar mapas visuais de assuntos;
- **Sistema de busca da BRAPCI** – faz uso da nuvem de tags como interface visual para auxiliar o usuário, sugerindo termos de busca.

5 ANÁLISE DOS DADOS E RESULTADOS

Nesta seção apresentam-se os resultados alcançados a partir da análise da relação entre OI e VI em *sites* de SRIs.

5.1 Sistemas que fazem uso da Visualização da Informação

Busca-se, a seguir, abordar alguns *sites* de SRIs que utilizam interfaces visuais em sua estrutura, demonstrando, de forma prática, como a aplicação de visualizações pode auxiliar na localização e recuperação de informação.


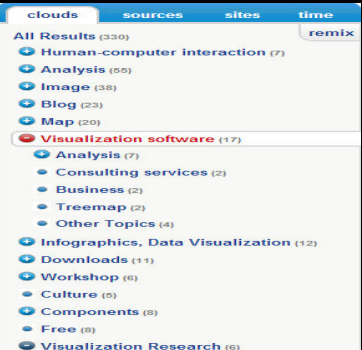
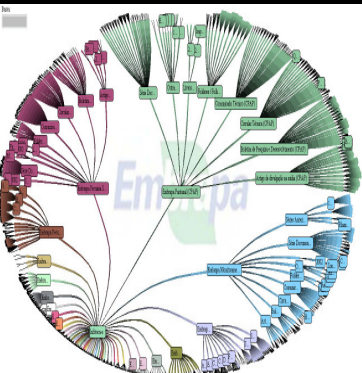
Nesse contexto os sistemas são analisados quanto as técnicas e modelos de interface de VI adotados em sua estrutura, levando em conta que, por meio de suas técnicas e modelos de interface, a VI pode ser considerada uma poderosa aliada para os sistemas que se propõem a oferecer uma interface amigável e interativa para acesso, localização e recuperação das informações por eles armazenadas.

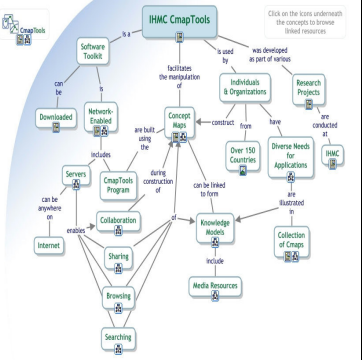
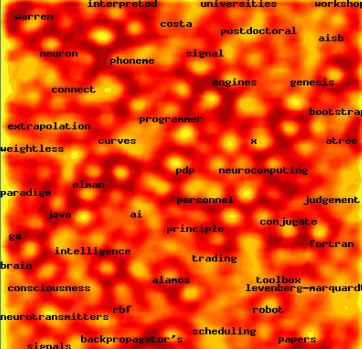

Por outro lado, os sistemas são analisados sob a perspectiva da organização da informação, ou seja, buscou-se compreender como deveriam estar organizadas as informações

nas bases de dados dos SRIs e quais instrumentos seriam necessários para tornar possível a construção de uma interface visual para esses sistemas, tendo em vista que a organização prévia das informações, bem como a adoção de instrumentos adequados para isso, são elementos necessários para a representação de informações por meio de estruturas de VI.

Nesse contexto, buscando demonstrar, de maneira resumida, os resultados obtidos por meio da análise realizada nos *sites* de SRI quanto à interação entre OI e VI, apresenta-se, a seguir, um quadro que permite visualizar as técnicas e modelos de interface visual adotados por cada um dos *sites*, bem como os instrumentos de OI adequados para a construção e funcionamento de visualizações expressivas.

QUADRO 1 – Relação entre OI e VI na aplicação de interface visual

Site de SRI	Técnica ou modelo de VI	Interface Visual	Instrumentos de OI para o emprego de visualização
1- <i>AquaBrowser</i> da biblioteca da Universidade de Edimburgo	Desenho de Grafos		Adoção de um tesouro ou taxonomia como instrumento de OI que permita uma representação e organização adequada de termos e que estabeleçam as relações existentes entre conceitos.
2- Mecanismo de busca YIPPY	Desenho de Árvores		Adoção de uma taxonomia como instrumento de OI que permita a categorização de conteúdos semelhantes, evidenciando as relações hierárquicas existentes entre eles.
3- Infoteca-e da Embrapa	Browser Hiperbólico		Adoção de uma taxonomia como instrumento capaz de organizar termos em categorias da mais genérica a mais específica, estabelecendo relações hierárquicas entre eles.

<p>4- Website do IHMC CmapTools</p> <p>Mapa Conceitual</p>		<p>Adoção de ontologia como instrumento que possibilite congregar conjuntos de conceitos padronizados, termos e definições, evidenciando a forma com que esses conceitos estão relacionados entre si.</p>
<p>5- WEBSOM - método para construção de mapas visuais.</p> <p>Cartografia Temática</p>		<p>Adoção da ontologia enquanto instrumento que permite a representação e a categorização de documentos, bem como o estabelecimento de temáticas relacionadas; adoção de critérios para identificar a coocorrência entre termos.</p>
<p>6- Sistema de busca da BRAPCI</p> <p>Nuvem de Tags</p>		<p>Adoção de listas de cabeçalhos de assunto para auxiliar na organização, estruturação e representação das tags que servirão como pontos de acesso para localização de documentos.</p>

Fonte: elaborado pelos autores

Assim, por meio do quadro 1, é possível visualizar os *sites* analisados, as técnicas e modelos de VI adotados por cada um, a figura ilustrando a interface visual, bem como os instrumentos de OI adequados para permitir uma eficaz organização e recuperação de informações. De maneira mais detalhada, tem-se:

Site 1- AquaBrowser da biblioteca da Universidade de Edimburgo - faz uso da técnica de VI “Desenho de Grafos” para disponibilizar o catálogo da biblioteca por meio de interface gráfica. A técnica Desenho de Grafos permite que, a partir de uma busca expressa na caixa de texto, o sistema gere uma visualização apresentando os resultados para essa busca por meio de termos, representados por vértices, e das relações entre eles, por arestas. Dessa forma, o termo utilizado na busca permanece no centro da figura enquanto os termos relacionados a ele estão ligados por meio das arestas, sendo o tipo de relação estabelecida

entre os termos determinada pela cor dos vértices, indicando associações, traduções, variações de ortografia e a trilha (rota) de pesquisa.

Quanto aos instrumentos de OI que podem ser adequados para organizar as informações no sistema de modo a evidenciar as relações existentes entre conceitos, tem-se o tesouro e a taxonomia. Os tesouros são capazes de organizar termos relacionados semântica e genericamente, e são considerados, conforme Carlan e Medeiros (2011), instrumentos de controle terminológico eficazes para a organização de conceitos e importantes ferramentas no tratamento e recuperação da informação.

No mesmo contexto está inserida a taxonomia, como instrumento que permite agrupamento categorizado de assuntos e a construção de relacionamentos semânticos entre termos, sendo considerada, também, como menciona Vital (2012, p. 26), “uma metodologia facilitadora tanto da organização quanto da recuperação, representação e disponibilização do conteúdo em ambiente Web”.

Site 2- Mecanismo de busca YIPPY - trata-se de um sistema de busca que faz uso de interface visual baseada na técnica “Desenho de Árvores” para organizar os itens recuperados, agregados em categorias e dispostos em uma lista ordenada, por meio de uma interface simples e intuitiva. A técnica Desenho de Árvores permite dispor conteúdos por hierarquias, por meio de uma árvore de diretórios, podendo disponibilizar itens por tópicos e subtópicos de acordo com a semelhança entre eles.

Nesse contexto, se faz necessário a análise de um instrumento de OI que auxilie na organização de conceitos, de maneira a possibilitar que a visualização cumpra o seu objetivo, ou seja, disponibilizar, de maneira organizada e simplificada, os resultados de uma busca. Sendo assim, pode-se citar a taxonomia como instrumento que busca estabelecer categorias para a informação e apresentam relacionamentos semânticos entre termos (VITAL, 2012). Assim, se torna possível apresentar os resultados de uma pesquisa por meio da categorização ou agrupamento de conteúdos semelhantes.

Site 3 - Infoteca-e da Embrapa - para fornecer acesso às informações sobre tecnologias produzidas pela Embrapa (Empresa Brasileira de Pesquisa Agropecuária) esse site adota uma árvore hiperbólica, isto é, uma interface visual em forma de árvore espacial, baseada na técnica “Browser Hiperbólico”. Essa interface permite que os conceitos estejam organizados por hierarquias (categorias), tendo um assunto mais geral centralizado e, dentro deste, assuntos específicos relacionados a ele, por meio de vértices em uma espécie de árvore.

Na estrutura, dentro de cada assunto específico estão dispostos os termos que se ligam a ele, por meio de ramificações que condensam os assuntos cada vez mais específicos, sendo

possível visualizar desde assuntos mais abrangentes até os menos inclusivos, possibilitando a compactação das informações à medida que se aproxima da periferia do desenho.

A eficaz utilização dessa técnica remete à necessidade de um instrumento de OI que permita uma organização baseada no estabelecimento de relações hierárquicas entre conceitos. Nesse contexto, se insere a taxonomia como instrumento voltado à organização de informações em ambientes específicos e que, como afirma Vital (2012), é caracterizado pela sua estrutura hierárquica, capaz de organizar termos em categorias da mais genérica a mais específica, estabelecendo relações entre conceitos.

Site 4 - Website do IHMC CampTools - apresenta um *software*, desenvolvido pelo IHMC (*Institute for Human & Machine Cognition*), que permite construir, navegar, compartilhar e criar modelos de conhecimento representados como mapas conceituais. Esse *website* disponibiliza um mapa para descrever a ferramenta por eles desenvolvida e para ilustrar o funcionamento de “Mapa Conceitual”, enquanto modelo de interface visual para organização e disponibilização de informações. Esse modelo de interface é formado por conceitos, geralmente localizados em círculos ou retângulos de algum tamanho, e por relacionamentos entre esses conceitos, que são indicados por uma linha que os conecta, formando proposições. Palavras sobre as linhas especificam as relações entre esses conceitos.

Para organizar as informações a serem dispostas por meio do mapa conceitual, pode-se citar como alternativa de instrumento de OI as ontologias, já que são conhecidas pela potencialidade em organizar e representar informação. As ontologias, como menciona Silva (2008), são estruturas que congregam conjuntos de conceitos padronizados, termos e definições, evidenciando a forma com que esses conceitos estão relacionados entre si. Nessa linha de pensamento, Brandão e Lucena (2002) enfatizam que o uso de ontologias possibilita não só a definição de conceitos num domínio de conhecimento, como também permite a definição de relacionamentos entre esses conceitos e de inferências básicas os envolvendo.

Site 5 - WEBSOM - método para construção de mapas visuais - é um método para construção de mapas visuais que organiza coleções de documentos e permite a exploração e recuperação desses documentos por meio de interface visual baseada na técnica “Cartografia Temática”. Essa técnica permite que os documentos armazenados estejam representados por meio de termos, que indicam o contexto dos documentos e ficam agrupados em pontos do mapa de acordo com a similaridade entre os conteúdos que representam, ou seja, a proximidade entre os termos indica que eles abordam temáticas relacionadas. Essa relação de associação entre os documentos é medida pela coocorrência entre termos. Dessa forma, os documentos semelhantes estão posicionados em grupos próximos uns dos outros no mapa.

Para que a técnica possa ser aplicada se faz necessário a adoção de um instrumento que permita representar os documentos por meio de um vocabulário controlado, que seja capaz de expressar as relações associativas entre os documentos. Sob essa perspectiva, pode-se citar como instrumento adequado ao contexto da cartografia temática as ontologias, tendo em vista que, como já mencionado por Silva (2008), as ontologias são estruturas que congregam conjuntos de conceitos padronizados, termos e definições aceitas por uma comunidade particular e a forma como esses conceitos estão relacionados entre si.

Site 6 - Sistema de busca da BRAPCI – a Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação adota visualizações em forma de “Nuvem de *Tags*” como uma alternativa amigável para a busca e identificação de conteúdos. Essa interface é composta por palavras-chave dispostas em uma lista visual com o objetivo de representar itens de conteúdo de um sistema. A nuvem de *tags* é, também, onde se concentram as *tags* mais utilizadas e/ou vistas no momento, podendo funcionar como pistas para os usuários, pois tem a função de facilitar a navegação, dando orientações e indicando por onde esses usuários poderiam caminhar para localização de informações de seu interesse.

Quanto ao instrumento a ser adotado para que esse tipo de interface possa ser utilizado, têm-se os cabeçalhos de assunto, tendo em vista que são instrumentos que uniformizam a descrição sintética dos assuntos contidos nos documentos, servindo como pontos de acesso para sua localização. Os cabeçalhos de assunto são elaborados a partir de palavras simples ou compostas (ordenadas em lista alfabética), geralmente extraídas do próprio documento, para descrever seus assuntos. Nas listas formadas pelos cabeçalhos de assunto podem ser encontradas, ainda, remissivas de orientação para assuntos relacionados.

As listas de cabeçalhos de assunto podem ser, portanto, úteis à organização da informação na nuvem de *tags*, visto que podem auxiliar na estruturação, representação e padronização das *tags*, bem como possibilitar a identificação e recuperação de documentos sobre uma determinada temática ou sobre temáticas relacionadas.

5.2 Relação entre OI e VI nos sites de SRIs que fazem uso de interface visual

Com base nas análises realizadas anteriormente acerca dos SRIs que adotam interfaces visuais em sua estrutura, é possível argumentar que, embora os sites de SRIs analisados apresentem propostas e funcionalidades diversificadas, todos eles caminham para os mesmos objetivos – facilitar a visualização, navegação, descoberta e recuperação de informações. Para alcançar esses objetivos, os sites adotam técnicas e modelos de interface visual que melhor se

adéquam aos seus contextos e particularidades. Em decorrência da adoção das interfaces visuais, esses *sites* oferecem diversos benefícios aos usuários, como:

- Oferecem uma maneira amigável e dinâmica de interação com as informações dispostas;
- Auxiliam na compreensão de como estão estruturadas as informações;
- Facilitam na localização de informações precisas;
- Contribuem para descoberta e dedução de novas informações,
- Permitem a visualização de conteúdos diferentes que antes eram desconhecidos, mas que podem vir a ser pertinentes à necessidade do usuário;
- Minimizam a sobrecarga de informação sobre o usuário.

Outro ponto importante a ser analisado quando se refere à aplicação de interfaces de VI em SRIs diz respeito aos aspectos que envolvem a organização prévia das informações a serem representadas.

A esse respeito, identificou-se a influência que a OI exerce, por meio de seus instrumentos, para auxiliar na organização e representação de conteúdos que serão transmitidos em estruturas visuais, confirmando, dessa forma, que sistemas que fazem uso de visualizações para transmitir informação requerem métodos de OI adequados para dar suporte às representações visuais de informação e para permitir a recuperação de conteúdos, já que, como afirma Zhang (2008), os métodos de organização e armazenamento de informações afetam e determinam os métodos de recuperação de informação.

Sob essa perspectiva, é importante elucidar que a atuação desses instrumentos de OI é direcionada para a organização das informações a serem transmitidas e que essa organização poderá ser refletida de maneira visual, podendo acarretar em visualizações expressivas, capazes de permitir representação facilitada, por meio de estrutura simples e recuperação satisfatória de informações.

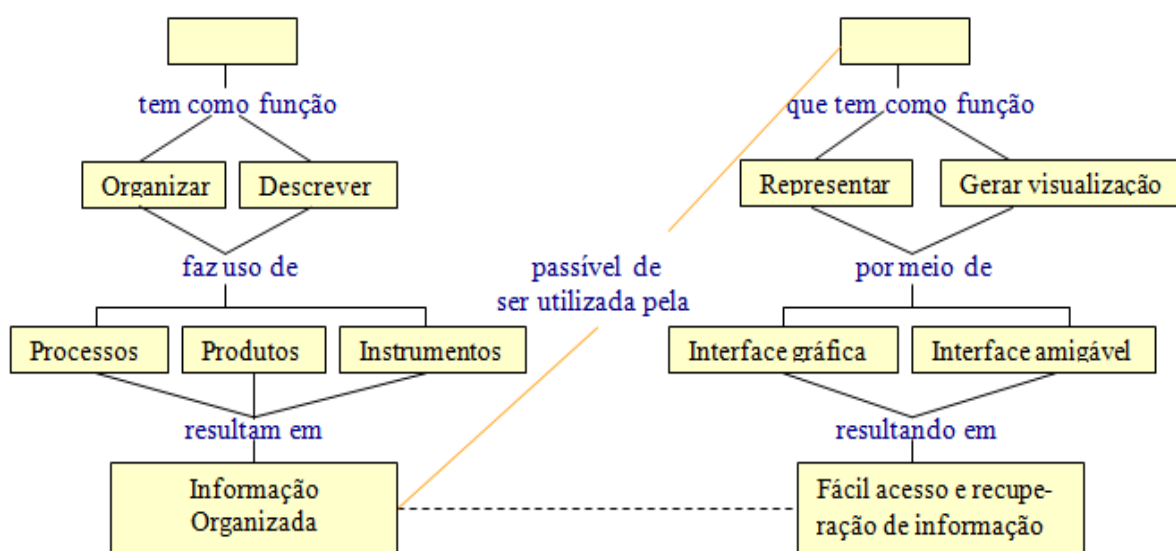
6 CONSIDERAÇÕES FINAIS

Os SRIs, enquanto sistemas que lidam com informação registrada e que trabalham com o propósito de permitir, a partir de um universo de documentos disponíveis, a recuperação daqueles que venham a ser úteis para satisfação de uma necessidade informacional, exercem um importante papel em uma sociedade caracterizada pela crescente produção e disseminação de informação, amplamente influenciada pelas tecnologias de informação e comunicação disponíveis.

A importante missão atribuída aos SRIs, a de tornar acessível uma quantidade cada vez maior de conteúdo informacional, implica em uma necessidade por meios precisos de acesso à informação – elemento vital às atividades humanas – bem como de maior controle sobre essa informação, com vista a prover a sua socialização e recuperação.

Considerando essas necessidades, partiu-se do pressuposto que a OI e a VI, enquanto áreas que assumem uma relação direta com a recuperação da informação, podem ser consideradas poderosas aliadas para os SRIs. Neste sentido, buscou-se uma análise mais aprofundada acerca dessas áreas, o que resultou em uma compreensão mais clara sobre as particularidades de cada uma e tornou possível identificar a potencial ligação entre elas. A FIG. 3 permite, por meio de uma visualização sucinta, ilustrar essa questão.

FIGURA 2 - Análise do foco das áreas de OI e VI



Fonte: elaborada pelos autores

A partir do que foi abordado na revisão de literatura acerca da OI e VI, foi possível evidenciar (conforme demonstrado na FIGURA 3) que a OI, enquanto área que se preocupa com a organização, descrição e tratamento da informação, fazendo uso, para isso, de processos, produtos e instrumentos, pode subsidiar a construção de interfaces de VI, cujo objetivo é representar informações de forma gráfica, facilitando o acesso e recuperação de conteúdos informacionais, por meio de uma interface amigável.

Dessa forma, entende-se que a OI é capaz de oferecer um suporte para que a VI exerça sua função de forma satisfatória, pois para que a VI potencialize o acesso e a recuperação da informação, o ideal é que haja uma prévia e adequada organização. Entende-se, assim, que a

OI e a VI possuem focos específicos, mas que podem caminhar em harmonia para um objetivo comum: possibilitar o acesso e a recuperação eficaz de informação organizada.

Foi exatamente nesse contexto que se desenvolveu o objetivo proposto por esta pesquisa, ou seja, buscou-se analisar a OI e a VI sob o ponto de vista da contribuição dessas áreas para os SRIs. Para isso, como validação do que foi pesquisado durante a revisão de literatura, foram analisados *sites* de SRIs que fazem uso de interfaces de VI, com o intuito de evidenciar as vantagens oferecidas pelo uso de visualizações nesses sistemas, bem como permitir a análise dos requisitos de OI necessários para tornar possível a construção das interfaces visuais e a recuperação de conteúdos informacionais relevantes.

Por meio da análise desses *sites* pode-se comprovar que, de fato, tem-se nas interfaces de VI ricas alternativas para representação de conteúdo informacional, disponibilizando uma maneira interativa para visualização e recuperação das informações armazenadas. Além disso, sob a perspectiva da OI, foi possível compreender a influência que a adoção de seus instrumentos exercem na organização e representação de conteúdos que serão transmitidos em estruturas visuais. Verificou-se que a adoção de instrumentos adequados pode resultar em uma recuperação eficaz de informações, corroborando, assim, que os recursos de OI afetam diretamente a recuperação da informação, conforme foi ressaltado ao longo da pesquisa.

Neste momento, é importante destacar que, embora não se possa afirmar que os *sites* analisados adotam em sua estrutura os instrumentos verificados como adequados nesta pesquisa, tais instrumentos foram sugeridos, pois entende-se que por meio da adoção deles se tem uma maior precisão na representação da informação, o que poderá possibilitar uma maneira expressiva de transmitir e recuperar informações por meio de interface visual, minimizando as chances de conteúdos relevantes permanecerem ocultos.

Em síntese, tem-se que as necessidades dos SRIs estão alinhadas à questão básica da CI, que gira em torno do acesso à informação, ou facilitação desse acesso. Assim, fundamentando-se no fato de que a CI está envolvida e empenhada nos aspectos que envolvem a recuperação da informação, bem como os meios de processá-la para otimizar sua acessibilidade e uso, considera-se, nesse cenário, a OI e a VI como potenciais auxiliadoras para os SRIs, considerando os preceitos de ambas as áreas e o papel que elas exercem quando se refere à informação, sua organização, visualização e recuperação.

Fundamentando-se na pesquisa desenvolvida, bem como nos autores referenciados neste trabalho chegou-se, por fim, à compreensão de que as áreas OI e VI convergem entre si, ou seja, partilham de objetivos comuns e, a partir disso, defende-se que essas áreas podem

alcançar promissores resultados, se atuarem em parceria para otimização na recuperação da memória armazenada em SRIs.

REFERÊNCIAS

- ALVARENGA, L. Representação do conhecimento na perspectiva da ciência da informação em tempo e espaço digitais. **Encontros Bibli: Revista eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v. 8, n. 15, p. 18-40, 1º sem. 2003.
- ARAÚJO, V. M. R. H. de. Sistemas de informação: nova abordagem teórico-conceitual. **Ciência da Informação**, Brasília, v. 24, n. 1, p. 1-39, jan./abr. 1995.
- BRANDÃO, A. A. F., LUCENA, C. J. P. de. **Uma introdução à engenharia de ontologias no contexto da web semântica**. Rio de Janeiro. Departamento de Informática, PUC-Rio, 2002. Disponível em: <ftp://ftp.inf.puc-rio.br/pub/docs/techreports/02_29_brandao.pdf>. Acesso em: 10 out. 2013.
- BRASCHER, M.; CAFÉ, L. Organização da informação ou organização do conhecimento?. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 9., 2008, São Paulo. **Anais...** São Paulo: ENANCIB, 2008. p. 1-14.
- BRASCHER, M.; CARLAN, E. Sistemas de organização do conhecimento: antigas e novas linguagens. In: ROBREDO, J.; BRÄSCHER, M. (Org.). **Passeios pelo bosque da informação: estudos sobre a representação e organização da informação e do conhecimento – eroic**. Brasília DF: IBICT, 2010. p. 148-176.
- CAFÉ, L.; SALES, R. Organização da informação: conceitos básicos e breve fundamentação teórica. In: ROBREDO, J.; BRÄSCHER, M. (Org.). **Passeios pelo bosque da informação: estudos sobre a representação e organização da informação e do conhecimento – eroic**. Brasília DF: IBICT, 2010. p. 115-129.
- CARLAN, E.; MEDEIROS, M. B. B. Sistemas de organização do conhecimento na visão da ciência da informação. **RICI: Revista Ibero-americana de Ciência da Informação**, Brasília, v. 4, n. 2, p. 53-73, ago./dez. 2011.
- DAL'EVEDOVE, P. R. A política de tratamento da informação na percepção de catalogadores de assunto. In: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA, DOCUMENTAÇÃO E CIÊNCIA DA INFORMAÇÃO, 24., 2011, Maceió. **Anais...** Maceió: CBBB, 2011. p. 1-12.
- DIAS, E. W. Contexto digital e tratamento da informação. **DataGramZero: Revista de Ciência da Informação**, Rio de Janeiro, v.2, n.5, p. 1-9, out. 2001.
- DIAS, M. P.; CARVALHO, J. O. F. de. A visualização da informação e a sua contribuição para a ciência da informação. **DataGramZero: Revista de Ciência da Informação**, Rio de Janeiro, v.8, n.5, p. 1-16, out. 2007.
- FREITAS, C. M. D.S. *et al.* Introdução à visualização de informações. **RITA: Revista de Informática Teórica e Aplicada**, Porto Alegre, v. 8, n. 2, p. 143-158, out. 2001.
- GIL, A. C. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Atlas, 2009.

GUIMARÃES, J. A. C. Abordagens teóricas de tratamento temático da informação (TTI): catalogação de assunto, indexação e análise documental. In: GARCÍA MARCO, F. J. Avances y perspectivas en sistemas de información y documentación. Ibersid, Zaragoza: Universidad de Zaragoza, 2009, p. 105-117.

LE COADIC, Y- F. **A ciência da informação**. 2. ed. Brasília: Briquet de Lemos/Livros, 2004.

MAZZA, R. **Introduction to information visualization**. New York, NY: Springer, 2009.

MCGARRY, K. **O contexto dinâmico da informação**: uma análise introdutória. Brasília, DF: Briquet de Lemos/Livros, 1999. 203 p.

NASCIMENTO, H. A. D.; FERREIRA, C. B. R. Visualização de informações: uma abordagem prática. In: CONGRESSO DA SOCIEDADE BRASILEIRA DE COMPUTAÇÃO, 25., 2005, São Leopoldo. **Anais...** São Leopoldo: UNISINOS, 2005. p. 1262 -1312.

SARACEVIC, T. Interdisciplinary nature of information science. **Ciência da Informação**, Brasília, v.24, n.1, p.36-41, jan./abr. 1995.

SILVA, D. L. da. **Uma proposta metodológica para construção de ontologias**: uma perspectiva interdisciplinar entre as ciências da informação e da computação. 2008. 286 f. Dissertação (Mestrado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2008.

SVENONIUS, E. The intellectual foundation of information organization. Cambridge: MIT Press, 2000.

VITAL, L. P. Taxonomia como elemento estruturante em portais corporativos. **Revista digital de Biblioteconomia e Ciência da Informação**, Campinas, v. 9, n. 2, p. 25-40, 2012.

ZHANG, J. Visualization for information retrieval. New York: Springer. 2008.

ANÁLISE DE ASSUNTO EM BIBLIOTECAS UNIVERSITÁRIAS: A OBSERVAÇÃO DE ESTRUTURAS TEXTUAIS EM LIVROS DAS ÁREAS DE HUMANAS, EXATAS E BIOLÓGICAS

SUBJECT ANALYSIS IN UNIVERSITY LIBRARIES: THE OBSERVATION OF TEXTUAL STRUCTURES IN HUMAN, EXACT AND BIOLOGICAL SCIENCES BOOKS

Daniela Majorie dos Reis
Mariângela Spotti Lopes Fujita

Resumo: A análise de assunto é um dos procedimentos realizados para o tratamento temático de livros em bibliotecas. Fatores como a observação de estruturas textuais dos documentos e do domínio em que o bibliotecário se encontra podem ser decisivos para a representação dos assuntos. A proposta desta pesquisa tem base em estudos sobre exteriorização de pensamentos especificamente a técnica de Protocolo Verbal. Foram coletados seis Protocolos Verbais Individuais, dois na área de Exatas, dois na área de Biológicas e dois na área de Humanas. O objetivo foi observar a análise de assunto nas três áreas do conhecimento e comparar as diferenças das estruturas textuais dos livros em cada área. As falas dos sujeitos foram analisadas e os resultados compilados em quadros. O quadro 3 apresenta sugestões para a adaptação dos modelos base, criando um modelo específico para o ensino da análise de assunto de livros científicos nas três áreas do conhecimento. O “Modelo de ensino de leitura documentária para análise de assunto de livros científicos” é uma fusão de dois modelos já existentes e tem ênfase no ensino da catalogação de assunto, o “Modelo de leitura documentária para indexação de artigos científicos” e o “Modelo de leitura documentária para indexação na catalogação de assuntos de livros em bibliotecas”.

Palavras-chave: Catalogação de Assunto. Protocolo Verbal. Biblioteca Universitária.

Abstract: The subject analysis is one of the procedures performed to the thematic treatment of books in libraries. Factors such as the observation of textual structures of the documents and the domain in which the librarian is located can be decisive for the representation of subjects. The purpose of this research has support on studies about externalization of thoughts, specifically on the technique of Verbal Protocol. Six individual verbal protocols were collected, two in the Exact Science area, two in the Biological Science area and two in the Human Science area. The purpose was to observe the subject analysis in the three areas of knowledge and compare the differences of textual structures of the books in each area. Some of the subjects speeches were analyzed and the results compiled in tables. Table 4 presents suggestions for adapting the base models, creating a specific model to teach subject analysis of scientific books in the three areas of knowledge. The “Modelo de ensino de leitura documentária para análise de assunto de livros científicos” is a fusion of two existing models and has as emphasis to teach subject cataloging. The base models are “Modelo de leitura documentária para indexação de artigos científicos” and “Modelo de leitura documentária para indexação na catalogação de assuntos de livros em bibliotecas”.

Keywords: Subject Cataloging. Verbal Protocol. University Library.

1 INTRODUÇÃO

Esta pesquisa surgiu da constatação da falta de uma metodologia de ensino específica que indique na estrutura textual as partes em que é possível identificar os conceitos para realizar a catalogação de assunto de livros científicos em bibliotecas universitárias.

Foi realizada investigação em leitura documentária de catalogadores para a exploração de estrutura textual na catalogação de assunto de livros com abordagem sociocognitiva em contexto de bibliotecas universitárias².

Buscou-se proporcionar elementos necessários para a elaboração de um “Modelo de ensino de leitura documentária para análise de assunto de livros científicos”, baseado no “Modelo de leitura documentária para indexação de artigos científicos”, e no “Modelo de leitura documentária para indexação na catalogação de assuntos de livros em bibliotecas”, a partir de resultados obtidos com levantamento bibliográfico e análise de dados coletados, ampliando o escopo de atuação da pesquisa.

O objetivo geral foi oferecer contribuições para a área de tratamento temático da informação a partir de investigações da catalogação de assunto visando a adaptação para o ensino da análise de assunto de dois modelos Fujita e Rubi (2006) e Fujita (2010), contribuindo para o aprimoramento de metodologias para catalogação de assunto de livros científicos em bibliotecas universitárias.

Definiu-se como objetivo específico, observar a análise de assunto de livros científicos em bibliotecas universitárias, nas áreas de Exatas, Humanas e Biológicas por meio de seis Protocolos Verbais Individuais (PVI).

A partir da análise dos PVIs, foi possível sugerir partes da estrutura textual consideradas importantes para o aprimoramento dos Modelos, além da validação das partes já incluídas nos mesmos.

Fujita e Rubi (2006) apresentam no artigo “Um modelo de leitura documentária para a indexação de artigos científicos: princípios de elaboração e uso para a formação de indexadores” o “**Modelo de Leitura Documentária para indexação de artigos científicos**” que relaciona as partes da estrutura do artigo (introdução, metodologia, resultados e conclusões) com conceitos que podem representa-lo (objetivo, ação, agente e etc) durante a indexação.

Para Fujita e Rubi (2006, p. 9) o uso do Modelo de Leitura Documentária:

² A proposta apresentada deriva de pesquisa de mestrado, iniciada em agosto de 2010 e com dissertação defendida em agosto de 2012 (REIS, 2012).

[...] dirige-se, primordialmente, para o ensino de indexação de graduação, pós-graduação e de profissionais que já atuam no mercado de trabalho. Neste sentido, o currículo do Curso de graduação em Biblioteconomia da UNESP – Campus de Marília integra a disciplina obrigatória “Leitura documentária” e no Programa de Pós-Graduação em Ciência da Informação da UNESP – Campus de Marília é oferecida a disciplina “Leitura Profissional” que utilizam o Modelo de Leitura associado aos conteúdos teóricos e metodológicos de seus planos de ensino.

O “Modelo de leitura documentária para indexação na catalogação de assuntos de livros em bibliotecas” apresentado por Fujita (2010) segue a mesma linha de pensamento do “Modelo de Leitura Documentária para indexação de artigos científicos”, porém, este tem como foco a análise de assuntos de livros em bibliotecas universitárias, como esta pesquisa.

Na estrutura do artigo são brevemente apresentados conceitos sobre indexação, catalogação de assunto, estrutura textual de livros e os procedimentos para aplicação do Protocolo Verbal, quais foram os critérios para escolha das bibliotecas pesquisadas e para a análise dos dados. Na sequência é exibida a análise dos dados coletados, onde algumas falas dos sujeitos foram selecionadas e comentadas. Os dados foram reunidos em quadros para melhor visualização. No quadro 3 são mostradas as sugestões para aprimoramento dos Modelos de Fujita e Rubi, 2006 e de Fujita, 2010. As partes mais consultadas por biblioteca e como estas partes afetam a análise de assunto foram comentadas. Finalizando as ideias que originaram a pesquisa, são feitas recomendações para o uso do Modelo adaptado, e como seu uso afeta a análise de assunto em bibliotecas universitárias e seu ensino.

2 PROCESSO DE INDEXAÇÃO E CATALOGAÇÃO DE ASSUNTO

Antes de apresentar algumas definições sobre indexação e catalogação de assunto, é importante ter em mente que um dos fatores que diferenciam ambas são seus produtos finais e o uso dos mesmos. A indexação é pensada em alguns momentos como processo voltado à elaboração de índices e a catalogação de assunto é o processo responsável pela seleção de descritores que serão inseridos em um catálogo junto à catalogação descritiva de documentos. Apesar das diferenças que seus produtos apresentam, ambas tem o processo de análise de assunto em comum.

Na visão das autoras Fujita, Rubi e Boccato (2009, p. 19):

[...] o processo de **indexação, catalogação de assunto**, classificação e elaboração de resumos, que são considerados processos de sumarização da informação dos quais se originam os **índices**, os **catálogos de assunto**, os números de classificação e os resumos que possibilitarão a recuperação da informação pertinente aos interesses dos usuários (grifo nosso).

Para Lancaster (2004, p. 6), a indexação de assuntos implica na preparação de uma representação do conteúdo temático dos documentos.

Chaumier (1988, p. 63) explica que a indexação é a parte mais importante da análise documentária, “conseqüentemente, é ela que condiciona o valor de um sistema documentário”. Seguindo a mesma linha de pensamento, para Fujita (2003, p. 62), a indexação é a parte mais importante em análise documentária por “condicionar os resultados de uma estratégia de busca”.

Guinchat e Menou (1994, p. 176) definem a indexação como uma das formas de descrição de conteúdo. É a operação pela qual se escolhe os termos mais apropriados para indicar o conteúdo de um documento.

Com relação à estrutura de documentos, Lancaster (2004, p. 24) afirma que:

As partes a serem lidas atentamente são as que apresentam maior probabilidade de dizer o máximo sobre o conteúdo no menor tempo: título, resumo, sinopse e conclusões. Os títulos das seções e as legendas das ilustrações ou tabelas também merecem maior atenção. Convém passar os olhos pelo restante do texto, para confirmar se as partes mais condensadas contêm uma imagem exata do que trata o documento.

Foskett (1973, p. 23) comenta que para realizar a análise de assunto “existem alguns atalhos que podemos tomar: ler o sumário, o prefácio ou a introdução, ou o comentário do editor na orelha da obra; talvez o autor haja preparado um resumo”.

Chaumier (1988, p.64) sugere que a leitura de um documento seja realizada dando mais atenção ao título, intertítulos, introdução, conclusão, frases introdutórias de parágrafos e capítulos, legendas de ilustrações, gráficos, tabelas, informações em negrito, entre outras, pois são as partes mais ricas em informação.

Além da observação de estruturas textuais, Cintra (1983, p. 6) explica que “a prática do indexador numa determinada área do conhecimento amplia seu glossário visual e torna o processo de indexação muito mais rápido”, o que justifica a delimitação de domínio para estudo nesta pesquisa.

Desta forma, é importante revelar a relação entre os acervos em assuntos específicos de bibliotecas universitárias, e como o tratamento destes documentos que a integram é afetado. O bibliotecário terá mais facilidade ao realizar a análise de assuntos de documentos cuja tematicidade lhe é familiar, daí a importância de seu conhecimento prévio adquirido durante os anos de experiência e suas estratégias de leitura documentária em determinada área do conhecimento.

Fujita (2004, p. 24) recomenda que para a formação do indexador em leitura documentária sejam considerados conhecimentos sobre: a estrutura textual, o processo de leitura documentária, o contexto sociocognitivo e físico dos serviços e os objetivos da leitura documentária baseada no conteúdo documentário e na demanda da comunidade usuária. Daí a

necessidade de observar a análise de assunto por meio do estudo de estruturas textuais específicas e do domínio em que se encontram, revelando como cada área e tipo de documento influenciam nos processos de análise de assunto.

3 METODOLOGIA

Para a observação da leitura documentária durante a tarefa de catalogação de assunto de livros científicos, foram selecionados seis bibliotecários catalogadores de bibliotecas de universidades públicas do estado de São Paulo nas três áreas do conhecimento (Humanas, Exatas e Biológicas) para coleta de dados a partir da técnica de Protocolo Verbal Individual (PVI).

A técnica introspectiva de coleta de dados escolhida para esta pesquisa, denominada “Protocolo Verbal Individual” será apresentada detalhadamente a seguir neste mesmo capítulo. Serão descritos os procedimentos que a permeiam, justificando porque foi a técnica escolhida para analisar o processo de catalogação de assunto nas bibliotecas selecionadas.

O ambiente universitário foi o escolhido, pois é praticamente “restrito” aos docentes e discentes da universidade interessados nos acervos especializados das bibliotecas da universidade. Dificilmente um usuário que busca informações mais gerais (como jornais, enciclopédias ou livros didáticos de nível médio) faria uso de uma biblioteca especializada. As bibliotecas universitárias brasileiras, segundo Fujita (2007, p.19) são consideradas “sistemas de informação que produzem bases de dados cujas formas de representação documentária estão organizadas em metadados com possibilidade de acesso múltiplo. São, portanto, instrumentos plurifuncionais porque dão acesso, confirmam dados e possibilitam avaliação”.

As bibliotecas universitárias foram escolhidas nas três áreas do conhecimento (Humanas, Exatas e Biológicas) com o intuito de comparar como o processo de análise de assunto é realizado em cada ambiente, ou seja, como as características e particularidades de cada área interferem na mente do bibliotecário no momento em que este realiza este processo.

Foram escolhidas bibliotecas públicas do estado de São Paulo, pois em todas é realizado obrigatoriamente o processo de catalogação de assunto, que é um dos requisitos organizacionais para as universidades participantes da rede CRUESP³ (algumas bibliotecas

³ O CRUESP/BIBLIOTECAS iniciou suas atividades em 1999, como Grupo de Estudos, instituído pela Resolução do Conselho de Reitores das Universidades Estaduais Paulistas (CRUESP) 149/99, tendo por objetivo a integração dos Sistemas de Bibliotecas da USP, UNESP e UNICAMP. Apresenta como missão: Consolidar o trabalho participativo e integrado dos

universitárias não realizam este processo, os termos são geralmente copiados de fichas catalográficas ou copiados de registros de outras bibliotecas).

A técnica escolhida foi o Protocolo Verbal Individual por revelar as estratégias metacognitivas dos bibliotecários com relação à catalogação de assunto. Vários autores contribuem com articulações relacionadas a técnicas introspectivas de coleta de dados, como K. Anders Ericsson em “*Protocol analysis and Verbal Reports on Thinking*”⁴. No capítulo seguinte serão apresentados conceitos sobre a técnica em questão.

Os Protocolos Verbais Individuais foram coletados durante a catalogação descritiva e temática de livros, e o principal ponto a ser analisado neste processo foi a parte de identificação e determinação de conceitos. Isto possibilitou a observação plena da identificação de conceitos por parte dos catalogadores, e as respectivas partes dos livros que foram consultadas para representar o documento.

Foi solicitado a cada bibliotecário que realizasse a catalogação de assunto de no mínimo três livros diferentes dentro da mesma área de assunto (Humanas, Biológicas ou Exatas), pois cada livro pode apresentar uma estrutura textual distinta, mesmo dentro da própria área.

Após a coleta e transcrição dos Protocolos Verbais Individuais, a análise dos dados foi realizada considerando o referencial teórico desenvolvido para esta pesquisa e com base em pesquisas anteriores que seguem a mesma linha de pensamento e que basearam o problema para esta pesquisa, como Sauperl (1999), Fagundes (2001), Rubi (2004), Silveira (2006) e Reis (2009). Foi dado maior enfoque às partes da estrutura textual de livros que são observadas com mais frequência no processo de catalogação de assunto.

A análise qualitativa dos PVIs proporcionou a observação do processo de catalogação de assunto de livros de áreas específicas em bibliotecas universitárias, ou seja, como determinada área interfere na análise de assunto de um catalogador em biblioteca universitária. Além da influência do ambiente universitário foi considerada principalmente a exploração da estrutura de livros científicos (entendidos aqui como livros cujos conteúdos são essencialmente científicos, criados a partir de hipóteses, relatos e fatos) durante sua análise.

Os PVIs foram aplicados apenas uma vez com cada bibliotecário tendo em vista fazer uso dos dados analisados para a criação de um “Modelo de ensino de leitura documentária

Sistemas das Universidades Estaduais Paulistas, buscando, principalmente, a cooperação, o compartilhamento e a racionalização dos recursos.

⁴ An updated and extracted version from Ericsson (2002).

para análise de assunto de livros científicos”, não sendo considerado necessário ou coerente reaplicar outro PVI com o uso do modelo elaborado com os bibliotecários que proporcionaram sua formulação.

3.1 A técnica do “pensar alto” – Protocolo Verbal

Existe um consenso de que a fala dos indivíduos na vida cotidiana reflete os aspectos de seus pensamentos, e normalmente as pessoas verbalizam apenas uma fração de todos os seus pensamentos (ERICSSON; SIMON, 1998, p. 179).

Pode-se observar então que nem tudo que pensamos é verbalizado durante o ato da fala, principalmente porque é considerado impossível uma pessoa verbalizar todo o fluxo de seus pensamentos, pois estes são processados pela mente de forma muito mais acelerada que a fala.

A possibilidade de saber o que cada pessoa está pensando por meio da exteriorização destes pensamentos durante algum processo pode ser útil por vários motivos. Um destes motivos são as pesquisas, estudos científicos sobre leitura, e nesta pesquisa, especificamente sobre leitura documentária.

Filósofos desde Aristóteles argumentaram que o pensamento é essencialmente um processo seqüencial, onde um pensamento leva a outro (ERICSSON; CRUTCHER, 1991 apud ERICSSON e SIMON, 1998). Desde então, as principais questões metodológicas têm se focado em determinar como obter informações sobre os estados de pensamento sem alterar a estrutura e a evolução das seqüências de pensamento que ocorrem naturalmente.

Para Fujita (2009, p. 51), a técnica de Protocolo Verbal:

[...] tem sido empregada como instrumento de pesquisa na coleta de dados que fornecem informações sobre processos mentais utilizados pelos indivíduos na realização de uma tarefa. É frequentemente usada em psicologia cognitiva e educação para observação e investigação dos processos mentais, especialmente em atividades de representação da informação e de uso de estratégias.

Ainda, na visão da autora, esta técnica “permite a observação do processo de leitura porque o leitor verbaliza o conhecimento processual que possui para o desenvolvimento da atividade” (FUJITA, 2009, p. 51).

Acredita-se que mesmo sob a “pressão” de estar constantemente falando alto durante uma gravação, e se isso não ocorre, ser incitado a “pensar alto”, a atividade não será influenciada pelo fato de que os sujeitos escolhidos possuem experiência de pelo menos cinco anos nas atividades praticadas (catalogação de assunto) e dificilmente se calarão diante de alguma dificuldade.

Para Ericsson e Simon (1998), a técnica do “pensar alto” tem ganhado aceitação como um método central e indispensável para estudos que tem como objetivo conhecer e entender melhor como o pensamento humano ocorre.

Assim, quando consideram-se os resultados dos protocolos verbais obtidos por meio do pensar alto, pode-se concluir que são fornecidas evidências detalhadas sobre as seqüências de pensamentos em uma ampla gama de tarefas, e as seqüências de pensamento verbalizado pelos participantes que realizam determinada tarefa são consistentes com as análises das tarefas, e os participantes, com o mesmo nível de habilidade, exibem características semelhantes de pensamento (ERICSSON; SIMON, 1998, p. 182).

Como mencionado anteriormente, a técnica de Protocolo Verbal foi escolhida para realizar a coleta de dados com os bibliotecários catalogadores por proporcionar a observação de suas estratégias metacognitivas durante a catalogação de assunto.

Estas estratégias de leitura devem ser estudadas e conhecidas, e o que se propõe nesta pesquisa é o estudo detalhado da exploração da estrutura textual de livros a partir dos relatos conscientes coletados nos PVIs.

Os procedimentos da coleta de dados com a técnica introspectiva de Protocolo Verbal são esquematizados em três momentos:

A) Procedimentos anteriores à coleta de dados:

- **Seleção do texto-base:** Deixou-se a critério do bibliotecário catalogador a escolha dos livros a serem analisados, lembrando que nenhum dos livros usados na coleta haviam sido tratados (catalogados) anteriormente pelo mesmo bibliotecário; foram selecionados livros específicos (no mínimo três) na área de cada biblioteca (Humanas, Exatas e Biológicas);
- **Seleção dos sujeitos:** Bibliotecários catalogadores responsáveis também pela catalogação de assunto (não só descritiva) de bibliotecas universitárias integrantes da rede CRUESP nas áreas de Humanas, Exatas e Biológicas;
- **Conversa informal com os sujeitos:** Foi solicitado aos catalogadores a realização da análise e catalogação de assunto de três ou mais livros em cada área, sempre “pensando alto”, revelando assim a exploração da estrutura textual dos documentos analisados;
- **Familiarização com a tarefa do “Think Aloud” (“Pensar Alto”) em sessões individuais:** Solicitou-se que cada catalogador realizasse a atividade como de costume, houve interação moderada por parte do pesquisador, somente quando necessário (quando o sujeito em análise ficava em silêncio).

B) Procedimentos durante a coleta de dados:

Gravação do “Pensar Alto” durante a leitura: Foram gravadas as exteriorizações do pensamento dos catalogadores durante a tarefa de catalogação de assunto de livros. Além da

gravação das falas dos sujeitos, foram feitas anotações pela pesquisadora com relação a expressões faciais e outros aspectos não capturados pela gravação, facilitando os procedimentos após a coleta de dados.

C) Procedimentos posteriores à coleta de dados:

- **Entrevista retrospectiva (opcional):** Foi realizada rápida entrevista retrospectiva com os sujeitos, onde se questionou quais partes da estrutura textual são mais usadas pelo mesmo durante o processo de análise de assunto além de outras perguntas consideradas pertinentes pela pesquisadora e/ ou que deixaram de ser comentadas pelo sujeito durante a coleta;
- **Transcrições literais das gravações:** As gravações das exteriorizações foram transcritas na íntegra utilizando algumas notações específicas para Protocolo Verbal adaptadas de Cavalcanti (1989)⁵.

Assim, após as coletas dos PVIs e de suas respectivas transcrições, foi criado o procedimento de análise das falas dos sujeitos, realizado com base na categoria “Alocamento de atenção a áreas importantes” de Brown (1980)⁶.

A seguir serão apresentadas as análises dos dados coletados e os resultados, onde foi realizado um exame detalhado de algumas falas de cada sujeito que faziam menção à exploração da estrutura textual dos livros analisados.

4 ANÁLISE DOS DADOS E RESULTADOS

A análise preliminar e detalhada das falas dos sujeitos foi realizada com base na categoria “Alocamento de atenção a áreas importantes” de Brown (1980). Foram observadas as principais partes consultadas por cada sujeito durante a análise do livro, e como cada uma dessas partes contribui para a catalogação de assunto.

Com base nas análises detalhadas das falas de cada um dos sujeitos, apresentadas a seguir e divididas por área do conhecimento (Humanas, Exatas e Biológicas respectivamente) foi elaborado um quadro apresentando uma síntese das partes consultadas na estrutura textual do livro em cada biblioteca (QUADRO 1).

ÁREA: HUMANAS: Biblioteca A

⁵ FUJITA, M. S. L., NARDI, M. I. A., FAGUNDES, S. A. A observação da leitura documentária por meio de protocolo verbal. In: RODRIGUES, G. M., LOPES, I. L. (Org.) **Organização e representação do conhecimento na perspectiva da ciência da Informação**. Brasília: Thesaurus, 2003. p. 141-178. (Estudos Avançados em Ciência da Informação, v.2)

⁶ BROWN, N. Metacognitive development and reading. In: SPIRO et al. (orgs). **Theoretical issues in reading comprehension**. New Jersey : L. Erlbaum Associate Publisers, 1980.

6 SUJEITO: A gente usa a página de rosto, a gente vem nessa ficha catalográfica que tem. [...] quando eu tenho muita dificuldade eu vou traduzindo entendeu? Eu traduzo aqui a parte de final do livro {mostra a contracapa do livro}, geralmente a gente usa também a conclusão. Da conclusão a gente tira muito assunto. [...] tem que olhar aqui dentro mesmo, aqui ó, isso aqui é **historia da arte**, ó, esse título aqui, esse aqui não tem ficha catalográfica, então o que a gente faz, será que é **filosofia**? {olha o sumário} Isso aqui vai ser “**filosofia**” porque fala de **dialética, identidade**, então a gente vai tirando tudo que tem. Mas aí a gente olha no meio, lê o resumo, a apresentação, a apresentação geralmente não porque a gente lê o sumário. E quando tem conclusão, a orelha.... então esse livro aqui vai ser **dialética**, dentro de **filosofia**. Então a gente dá uma olhada {consulta os capítulos do livro} [...].

Na Biblioteca A o sujeito consultou a folha de rosto, ficha catalográfica, contracapa do livro, conclusão, título, sumário, resumo, orelha e os capítulos do livro. Na fala 18, ao observar o título de um livro o sujeito pensa que o assunto principal será “número”, pois o título era “Cogitações sobre o número”. Ao consultar outras partes do livro, chega à conclusão de que não será “número” e sim “arte” o assunto principal.

ÁREA: HUMANAS: Biblioteca B

4 SUJEITO: [...] eu tinha visto que era sobre **estética na Idade Média**, pensando no título [Estética na Idade Média], ele me dá alguma indicação. Aí, eu tava olhando mesmo aqui o sumário, a catalogação na fonte, que ajuda, né? [...] Eu vou ver se tem mais alguma coisa interessante aqui no sumário {examina o sumário com cautela}... Agora eu vou acrescentar umas coisas que eu encontrei aqui no sumário e no título mesmo. [...] Às vezes eu também olho a orelha do livro, a contracapa, porque às vezes tem algum comentário, alguma coisa importante [...] pelo título [Filosofia e música] parece que tem a ver. Aí também tem a catalogação na fonte, o sumário [...], aí eu achei uma outra coisa no sumário que pode ser interessante pra gente que é/ fala de “**notação musical**”, que eu vou acrescentar também [...].

Na Biblioteca B, o sujeito consultou o título, sumário, catalogação na fonte, orelha, contracapa e apresentação do livro. No final da fala 14 o sujeito explica que em alguns livros, o título pode ser bem autoexplicativo, mas requer pesquisa externa ao seu conteúdo para ser eficaz (Funambulismo = a conhecida arte circense da “corda-bamba”). Neste caso o título foi de grande importância na determinação dos assuntos do livro.

Nas Bibliotecas A e do B a dificuldade para encontrar o assunto na estrutura textual pode ser grande independente do livro ou outro tipo de documento como partituras ou catálogos de arte. Os bibliotecários encarregados da análise de assunto possuem um esquema de busca que proporciona melhores resultados com relação à complexa estrutura dos documentos da área de Humanas, mesmo que suas estratégias envolvam a leitura de quase ou todas as partes da estrutura textual do livro. O título pode ser uma parte importante a ser observada para a análise de assunto, mas como foi visto nas duas bibliotecas, quando os livros são da área de Humanas, é importante considerar outras fontes em conjunto com o título dos livros.

ÁREA: EXATAS: Biblioteca C

4 SUJEITO: [...] Então o assunto dele, as palavras chave {consultando o título do livro} é linguagem de programação c++ e computer graphic, que é computação gráfica [...] Vejo o contents {sumário}, vou lendo pelo contents, que os livros são todos em inglês, francês, muito em alemão, aí, por exemplo, quando tem o problema da linguagem, porque tem até russo. [...] Então eu faço assim, vou pelo contents, daí pelo contents [...] quando você pensa já entra no texto, e os textos são tudo quebradinho. [...] Aí, uma coisa que eu gosto muito de olhar é o prefácio. Porque a parte de prefácio, ela sempre vai me dar informação do livro [...] o abstract, lembra isso é importante porque ele sempre traz alguma informação sobre livros da área dele, de alguma coisa que tenha familiarizado ao livro. [...] E o contents. [...] Eu dou uma olhada pelo título, porque como é a parte de Exatas, a parte de Exatas é bem direcionada, pelo, focada no título. Pelo tipo de livro que a gente tem, é bem assim mesmo. Porque o título da parte na área de Exatas, o título é bem o referencial. Não é como na parte de Humanas que no título ele vai tá contando alguma história, né? Aqui o título é mais exato. Mas nem sempre também a gente vai pelo título. ((RI)) [...] existem casos também que você não pode ir pelo título porque o texto do contents localiza outra informação [...].

Na Biblioteca C o sujeito consultou o título, o sumário, prefácio e o *abstract* (resumo da obra). De acordo com o sujeito, por ser uma biblioteca da área de Exatas, o assunto dos livros é bem representado em seus títulos, sendo necessária apenas uma conferência nas outras partes mais ricas em informação, que no caso destes livros, na opinião do sujeito, é o sumário e o prefácio.

ÁREA: EXATAS: Biblioteca D

2 SUJEITO: [...] Pelo título, eu já vejo mais ou menos onde ele vai se encaixar. Por exemplo, se tá falando de “administração de *gigabytes*”, é de computação. [...] Porque sem você souber o que tá escrito no título não dá nem pra você começar a fazer nada. [...] Então eu acho que, ah, tem uma coisa também que traz a informação do assunto, é a ficha catalográfica, tem uns livros que tem a ficha catalográfica aqui atrás {refere-se ao verso da folha de rosto} [...] aí eu já deixo anotado aqui no papel o termo que vai ser, e o outro é o ((FR)) próprio título do livro quase, você percebeu? Então, aqui, o primeiro livro que eu fiz, eu tirei os dados do título, o assuntos, e um pouco do conteúdo dele, na verdade o sumário. E desse livro aqui eu tirei da própria ficha catalográfica que tava atrás dele da *Lybrary*, na pagina de rosto. [...] Então esse daqui já tá definido e eu me baseei pela ficha catalográfica que tá no verso da folha de rosto, e também no título, porque sem eu ver a ficha eu já tinha definido. [...] Prefácio, no prefácio também é bom a gente dar uma lida, uma pesquisada pra saber. [...] Ó, aqui na introdução, na introdução tá mais bacana [...].

Na Biblioteca D o sujeito consultou o título, o sumário, a ficha catalográfica, o prefácio e a introdução. O sujeito usou o título dos livros para basear sua ideia de assunto principal do livro. Por não acreditar que o título seja confiável em todas as situações, o sujeito também retira informações do sumário e da ficha catalográfica.

Nas Bibliotecas C e D, os sujeitos fizeram bastante uso das informações contidas no título e no sumário dos livros. Os livros da área de Exatas, diferente dos livros da área de Humanas, apresentam uma estrutura mais uniforme de maneira geral, com um título bem representativo e um sumário que lista de forma completa todos os conteúdos apresentados nos

capítulos do livro. Esses fatores auxiliam o bibliotecário no momento da análise de assunto, mesmo quando os livros estão em outra língua.

ÁREA: BIOLÓGICAS: Biblioteca E

4 SUJEITO: [...] a gente analisa o título, e o folha de rosto, ficha catalográfica, e introdução e sumário? E às vezes quando tem a orelha? A gente avalia, ou a contracapa também, dependendo se tem alguma informação, a gente dá uma avaliada do conteúdo. [...] Esse livro aqui por exemplo [Ortodontia] {leitura do título} com o Moyers {autor do livro} aqui, é assim a gente lá, lá a gente/ o dia-a-dia, já é conhecido, é uma obra bem conhecida [...] então, a indexação acaba sendo mais genérica, inclusive dessa livro aqui [...] a gente faz até uma avaliação do prefácio, do que a obra se trata. E deixa eu ver aqui, vamos fazer a leitura do prefácio [...] Nesse caso aqui, além da Ortodontia vai lendo aqui o prefácio e algumas partes do sumário [...] Agora com relação a esta outra obra [...] vai ver a folha de rosto, como nessa outra também não tem/ o título é [Ortodontia preventiva básica] {leitura do título} do [Alael de Paiva Lima] {leitura do autor do livro} a gente vê a folha de rosto, daí dá uma olhada no/ faz uma leitura do prefácio... Como também na contracapa pra ver se tem alguma coisa, e na/ no prefácio tem um prólogo aqui, que eu vou dar uma lida pra ver do que se trata, em termos de assunto... {leitura breve do prólogo} vou ver o sumário também {consulta breve do sumário} [...].

Na Biblioteca E o sujeito consultou o título, folha de rosto, ficha catalográfica, introdução, sumário, orelha, contracapa, prefácio e prólogo. O sujeito explica a importância do reconhecimento do autor no momento da análise de assunto de um livro, pois um autor pode escrever sobre determinado assunto especificamente, e sabendo disso, o bibliotecário saberia qual o assunto principal de um livro só por conhecer seu autor.

ÁREA: BIOLÓGICAS: Biblioteca F

2 SUJEITO: [...] Então o primeiro passo, principalmente você vai nas partes principais do livro, folha de rosto, eu sempre olho a folha de rosto, gosto de olhar aqui a última capa, porque a gente tem uma boa idéia de crítica, e muito do assunto também. [...] só de olhar pelo título você já sabe né mais ou menos qual desses cinco assuntos você vai encaixar. Aí nesse aqui eu vou começar pela página de rosto, pra leitura técnica que eu faço eu gosto de ler todo o conteúdo {refere-se ao contents do livro}, leio o conteúdo, leio o prefácio, porque por mais que/ por exemplo, o título seja bem explícito, né? [...] Só com o título você não consegue ter essa especificidade, esse aqui é sobre “proteínas”, “evolução”, tá vendo? [...] Então o conteúdo, você lê os contents, o prefácio, lê muitas vezes a introdução, é muito importante. [...] Eu usei a folha de rosto, os contents, aqui também, eu uso bastante essa fichinha, a fichinha catalográfica na fonte, me baseio muito por ela também. [...] Bom, esse daqui tem um contents bem/ tá bem rico [...].

Na Biblioteca F o sujeito consultou a folha de rosto, contracapa, título, sumário, prefácio, introdução, ficha catalográfica e contracapa. O sujeito explica que a primeira parte da estrutura que observa é o título, pois é a partir dele que definirá o assunto principal do livro dentro das subáreas da Biologia. Explica que consulta as partes citadas em ordem como rotina e não como conferência dos assuntos retirados do título.

Nas Bibliotecas E e F, a rotina de leitura para a análise de assunto é semelhante à das bibliotecas da área de Exatas. Os sujeitos da área de Biológicas observaram partes em comum

aos sujeitos da área de Exatas, e deram mais atenção ao título também, porém consultaram mais partes para conferir os assuntos retirados do título dos livros.

Pode-se observar que os Protocolos Verbais Individuais foram coletados durante a catalogação de assunto de livros específicos nas áreas das bibliotecas universitárias. As bibliotecas selecionadas são de universidades públicas do estado de São Paulo nas três áreas do conhecimento integrantes da rede CRUESP: Bibliotecas A, B, C, D, E e F.

A partir da análise detalhada das falas foi elaborado o quadro 2, que apresenta uma comparação entre as partes da estrutura textual de livros sugeridas na literatura e as partes equivalentes consultadas pelos catalogadores das bibliotecas analisadas (QUADRO 1):

QUADRO 1 – Partes consultadas na estrutura textual do livro durante a análise de **assunto**

PARTES CONSULTADAS NA ESTRUTURA TEXTUAL DO LIVRO DURANTE A ANÁLISE DE ASSUNTO					
BIBLIOTECA A	BIBLIOTECA B	BIBLIOTECA C	BIBLIOTECA D	BIBLIOTECA E	BIBLIOTECA F
Título; Folha de rosto; Ficha catalográfica; Contracapa; Conclusão; Resumo; Apresentação; Orelha; Capítulos do livro (desenvolvimento).	Título; Subtítulo; Sumário; Catalogação na fonte; Orelha do livro; Contracapa; Folha de rosto; Introdução; Apresentação.	Título; Sumário; Prefácio; Resumo/ Abstract;	Título; Subtítulo; Autor (campo de pesquisa do autor); Sumário; Prefácio; Introdução	Título; Folha de rosto; Ficha catalográfica; Introdução; Sumário; Orelha do livro; Contracapa; Prefácio; Capa; Biografia do autor;	Título; Folha de rosto; Catalogação na fonte; Sumário; Prefácio; Introdução; Contracapa;

Fonte: elaborado pela autora.

QUADRO 2 - Comparação das sugestões de estrutura textual por autores e das partes consultadas durante a análise de assunto pelas bibliotecas pesquisadas

Partes do livro: NBR 12676	Autores da Ciência da Informação					Autores de outras áreas	Bibliotecas pesquisadas
	Foskett (1973)	Langridge (1977)	Chaumier (1988)	Lancaster (2004)	Taylor (2004)	Kato (1995)	
Título e subtítulo	---	Título	Título e intertítulos	Título	Título e subtítulo	---	Todas.
Resumo se houver	Resumo	---	---	Resumo	Resumo	Resumo	Biblioteca A; Biblioteca C.

Sumário e equivalentes	Sumário	Listas de conteúdo	---	---	Tabela de conteúdos/sumário	---	Biblioteca B; Biblioteca C; Biblioteca D; Biblioteca E; Biblioteca F.
Introdução	Prefácio e/ou introdução	Introduções	Introdução	---	---	---	Todas.
Palavras em destaque	---	Cabeçalhos dos capítulos	Frases introdutórias de parágrafos e capítulos, informações em negrito	Títulos das seções	---	Capítulos / Seções	Biblioteca A.
Ilustrações, tabelas e diagramas	---	---	Legendas das ilustrações, gráficos, tabelas	Legendas das ilustrações e/ou tabelas	Ilustrações, diagramas, tabelas e legendas	---	Nenhuma.
Partes diversas (incluindo referências bibliográficas)	Comentário do editor na orelha da obra (Elemento opcional pela NBR 6029)	---	Conclusão	Sinopse e conclusões	Palavras, ou frases impressas em fontes diferentes do resto do texto; hyperlinks	---	Biblioteca A; Biblioteca B; Biblioteca E.

Fonte: Elaborado pela autora.

A partir do QUADRO 2 foi possível observar que o **Título** e a **Introdução e/ ou Prefácio** foram as partes consultadas por todos os catalogadores. O **Título** que embora não seja recomendado para a catalogação de assunto por vários motivos, provou ser bastante útil em todas as áreas do conhecimento (inclusive na área de Humanas) quando usado em conjunto com outras partes como a **Introdução e/ ou Prefácio** e **Sumário e equivalentes**. Ambas as partes da estrutura foram sugeridas para o modelo.

O **Sumário e equivalentes** foram a terceira parte da estrutura textual mais consultada, não sendo observada apenas pelo catalogador da Biblioteca A. Isso mostra que o **Sumário e equivalentes** é uma parte da estrutura que deve ser considerada importante na catalogação de assunto em bibliotecas universitárias.

O **Resumo** foi consultado apenas pelo catalogador da Biblioteca A (Humanas) e pelo catalogador da Biblioteca C (Exatas). Considerando o fato de que nem todos os livros apresentam um resumo em sua estrutura textual, esta foi uma parte não selecionada para compor o modelo.

Os **Títulos dos capítulos e/ou capítulos** e as **Conclusões** foram partes consultadas apenas pelo catalogador da Biblioteca A. Acredita-se que este catalogador especificamente tenha consultado partes diferentes de todos os outros sujeitos devido ao fato dos livros estarem em outras línguas que não o inglês, espanhol ou português, dificultando sua análise de assunto. Também, a maioria dos livros catalogados pelo sujeito eram antigos e com uma estrutura completamente diferente dos livros atuais (muitos não apresentavam folha de rosto e dados sobre a publicação em sua estrutura).

As partes diversas consultadas pelos catalogadores da Biblioteca A, B e E são o Comentário do editor na orelha da obra (Elemento opcional pela NBR 6029) e as Palavras, ou frases impressas em fontes diferentes do resto do texto; hyperlinks. Estas são partes que embora tenham sido pouco consultadas pelos catalogadores, são de rápida absorção e geralmente são informações curtas e significativas, principalmente para a área de Humanas (dois dos três catalogadores que consultaram as partes), porém, não foram incluídas no Modelo.

As **Ilustrações, gráficos e diagramas** são partes que não foram consultadas por nenhum dos catalogadores (inclusive os catalogadores que trabalham com catálogos de arte e livros de ilustradores), e portanto não foram sugeridas no Modelo proposto a seguir.

QUADRO 3 - Modelo de ensino de leitura documentária para análise de assunto de livros científicos

CONCEITOS	QUESTIONAMENTO PARA IDENTIFICAÇÃO DE CONCEITOS	PARTES DA ESTRUTURA DO LIVRO
OBJETO e PARTE(S) DO OBJETO (algo ou alguém que está sob estudo do autor)	O documento possui em seu contexto um objeto sob efeito desta ação?	TÍTULO, SUBTÍTULO, SUMÁRIO E INTRODUÇÃO (nesta ordem)
AÇÃO (processo sofrido por algo ou alguém)	O assunto contém uma ação (podendo significar uma operação, um processo etc.)?	TÍTULO, SUBTÍTULO, SUMÁRIO E INTRODUÇÃO (nesta ordem)
AGENTE (aquele ou algo que realizou a ação)	O documento possui um agente que praticou esta ação?	TÍTULO, SUBTÍTULO, SUMÁRIO E INTRODUÇÃO (nesta ordem)
MÉTODO (métodos utilizados para realização da pesquisa)	Para estudo do objeto ou implementação da ação, o documento cita e/ou descreve modos específicos, por exemplo: instrumentos especiais, técnicas, métodos, materiais e equipamentos?	SUMÁRIO INTRODUÇÃO OU METODOLOGIA
TEMPO (ano, período ou época)	O estudo foi desenvolvido em período específico? É relevante representá-los na Catalogação de Assunto?	SUMÁRIO INTRODUÇÃO OU METODOLOGIA
LOCAL OU AMBIÊNCIA (local físico onde foi realizada a pesquisa)	Todos estes fatores são considerados no contexto de um lugar específico ou ambiente?	SUMÁRIO INTRODUÇÃO OU METODOLOGIA

PONTO DE VISTA DO AUTOR [perspectiva, opinião]	O assunto foi considerado de um ponto de vista, normalmente não associado com o campo de estudo (por exemplo, um estudo sociológico ou religioso)?	INTRODUÇÃO E PREFÁCIO
CAUSA E EFEITO Causa (ação+objeto)/Efeito	Considerando que a ação e o objeto identificam uma causa, qual é o efeito desta causa?	CONCLUSÕES OU PARTE FINAL

Fonte: Elaborado pela autora com base em Fujita (2010).

De acordo com as instruções de como utilizar o “Modelo de leitura documentária para indexação na catalogação de assuntos de livros em bibliotecas⁷” os três conceitos **objeto**, **ação** e **agente** (os três primeiros no quadro 3) são principais, estão interligados e são dependentes um do outro e devem ser identificados na estrutura textual em ordem.

Portanto, foi nos três conceitos principais que considerou-se importante fazer sugestões. As partes acrescentadas estão em realce cinza nas linhas correspondentes aos conceitos: objeto, ação e agente.

O **Sumário e equivalentes** foi uma das partes selecionadas na análise dos dados para compor o modelo, mas esta é uma parte da estrutura textual que já faz parte do Modelo de leitura documentária para indexação na catalogação de assuntos de livros em bibliotecas, então, não foi acrescentada.

Outra parte importante observada pelos sujeitos foi a **Introdução e/ ou Prefácio** que também foi selecionada para compor o Modelo de ensino de leitura documentária para análise de assunto de livros científicos, mas assim como o Sumário e equivalentes não foi acrescentada, pois já fazia parte do modelo que serviu de base.

O **Título** foi praticamente a única parte escolhida para compor o modelo, pois foi consultado por todos os sujeitos e mostrou-se bastante útil, mesmo para os catalogadores da área de Humanas. Na análise dos dados, as falas dos sujeitos mostram sua postura quanto ao uso do título para a análise de assunto. Retomando algumas das falas analisadas:

- Na Biblioteca A o sujeito ao observar o título de um livro pensa que o assunto principal será “número”, pois o título era “Cogitações sobre o número”. Ao consultar outras partes do livro, chega à conclusão de que não será “número” e sim “arte” o assunto principal;
- Na Biblioteca B, o sujeito ao consultar o título explica que em alguns livros, o título pode ser bem autoexplicativo, mas requer pesquisa externa ao seu conteúdo para ser eficaz (Funambulismo = a conhecida arte circense da “corda-bamba”). Neste caso o título foi de grande importância na determinação dos assuntos do livro;

⁷ Em Fujita e Rubi (2006).

- Na Biblioteca D o sujeito consultou o título e por não acreditar que o título seja confiável em todas as situações, também retira informações do sumário e da ficha catalográfica;
- Na Biblioteca F o sujeito explica que a primeira parte da estrutura que observa é o título, pois é a partir dele que definirá o assunto principal do livro dentro das subáreas da Biologia. Explica que consulta as partes citadas em ordem como rotina e não como conferência dos assuntos retirados do título.

Portanto, a única parte consultada por todos os sujeitos durante a análise de assunto que pôde ser acrescentada ao Modelo de Fujita (2010) foi o **Título e subtítulo** dos livros, pois todas as outras partes consultadas já faziam parte do mesmo.

Mesmo que a análise dos dados resultou na sugestão de apenas uma parte da estrutura textual nova para o Modelo, considera-se importante todo o trabalho de pesquisa envolvido. As partes da estrutura analisadas provam que o Modelo de leitura documentária para indexação na catalogação de assuntos de livros em bibliotecas de Fujita (2010) é muito eficaz e condiz com a catalogação de assunto na realidade das bibliotecas universitárias nas três áreas do conhecimento.

5 CONCLUSÃO

A partir da análise das falas dos sujeitos, foi possível observar as partes da estrutura textual mais usadas para a análise de assunto dos livros em cada biblioteca. Os resultados desta análise detalhada são apresentados no quadro 1.

A análise detalhada que resultou no quadro 1 proporcionou a observação da relevância do uso de cada parte do documento para o processo de análise de assunto em bibliotecas universitárias, evidenciando a necessidade de sua inclusão como parte sugerida para consulta no “Modelo de ensino de leitura documentária para análise de assunto de livros científicos” apresentado no QUADRO 3.

A elaboração do Modelo proposto foi realizada com base na estrutura do modelo de Fujita e Rubi (2006), Fujita (2010) e na análise aprofundada das falas dos catalogadores sobre o uso das partes da estrutura textual de livros na síntese apresentada no quadro 1, e no quadro 2 apresentados nos resultados, levando em consideração também o que autores da área sugerem para leitura em análise de assunto.

O Modelo de ensino de leitura documentária para análise de assunto de livros científicos (QUADRO 3) foi cuidadosamente elaborado pensando nas partes da estrutura textual já sugeridas para a catalogação de assunto no Modelo base. Isso resultou na sugestão

de apenas uma parte para o Modelo, incluídas nos três conceitos principais (Objeto, ação e agente), o título dos livros.

Os resultados apresentados reafirmam a importância dos Modelos usados como base para o “Modelo de ensino de leitura documentária para análise de assunto de livros científicos” ao validar as partes da estrutura textual sugeridas pelos mesmos.

Recomenda-se que o Modelo de ensino de leitura documentária para análise de assunto de livros científicos seja usado em cursos de graduação. Ao preparar discentes de cursos de biblioteconomia para uma análise de assunto feita com base na estrutura de documentos, ocorrerá uma conscientização de que este é um processo importante na recuperação pelo usuário (independente dos seus produtos finais, catálogos ou índices), preparando-os desde cedo para uma das práticas profissionais mais importantes da área.

É imperativo que discentes de cursos de biblioteconomia entendam que o tipo de biblioteca (pública, escolar, especializada ou universitária) e seus respectivos usuários definirão como será feito o tratamento temático dos documentos, e por este motivo não se deve em hipótese alguma simplesmente copiar registros importados de outras bibliotecas junto com a catalogação descritiva, pois isso empobrece a representação do documento e impossibilita a recuperação pelo usuário.

REFERÊNCIAS

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS - ABNT. **NBR 12676. Métodos para análise de documentos:** determinação de seus assuntos e seleção de termos de indexação. Rio de Janeiro, 1992.

CHAUMIER, Jacques. (1988). **Indexação:** conceito, etapas, instrumentos. Trad.

CINTRA, Ana Maria Marques. Elementos de lingüística para estudos de indexação. **Ciência da Informação**, v.12, n.1, p.5-22, 1983.

ERICSSON, K. Anders; SIMON, Herbert A. How to Study Thinking in Everyday Life: Contrasting Think-Aloud Protocols With Descriptions and Explanations of Thinking. **Mind, Culture, And Activity**, Califórnia, v. 3, n. 5, p.178-186, 1998.

FAGUNDES, Silvana Aparecida. **Leitura em análise documentária de artigos de jornais.** 2001. 322 f. Dissertação (Mestrado) - Departamento de Ciência da Informação, Unesp, Marília, 2001.

FOSKETT, Antony Charles. **A abordagem temática da informação.** Trad. Antônio Agenor Briquet de Lemos. São Paulo: Polígono; Brasília: UnB, 1973.

FUJITA, M. S. L. O contexto da leitura documentária de indexadores de bibliotecas universitárias em perspectiva sócio-cognitiva para a investigação de estratégias de ensino. Marília: Faculdade de Filosofia e Ciências, UNESP, 2010. 115f. (Relatório Final de Pesquisa, Bolsa CNPq de Produtividade em Pesquisa – Nível 2)

FUJITA, Mariângela Spotti Lopes (Org.). **A indexação de livros: a percepção de catalogadores e usuários de bibliotecas universitárias.** São Paulo: Cultura Acadêmica, 2009. v. 1. 104 p.

FUJITA, Mariângela Spotti Lopes, NARDI, Maria Isabel Asperti, FAGUNDES, Silvana Aparecida. A observação da leitura documentária por meio de protocolo verbal. In: RODRIGUES, G. M., LOPES, I. L. (Org.) **Organização e representação do conhecimento na perspectiva da ciência da Informação.** Brasília: Thesaurus, 2003. p. 141-178. (Estudos Avançados em Ciência da Informação, v.2)

FUJITA, Mariângela Spotti Lopes, RUBI, Milena Polsinelli. Um modelo de leitura documentária para a indexação de artigos científicos: princípios de elaboração e uso para a formação de indexadores. **Datagramazero**, Rio de Janeiro, v. 7, n. 3, p.1-18, 2006.

FUJITA, Mariângela Spotti Lopes. A abordagem cognitiva da leitura como prática pedagógica no ensino da disciplina leitura documentária no curso de Biblioteconomia da UNESP - Campus de Marília: uso do protocolo verbal para metacognição do indexador aprendiz. In: SANTOS, Jussara Pereira Santos (Org.). **A leitura com prática pedagógica na formação do profissional da informação.** Rio de Janeiro: Fundação Biblioteca Nacional, 2007. p. 101-132.

FUJITA, Mariângela Spotti Lopes. A leitura documentária na perspectiva de suas variáveis: leitor-texto-contexto. **Datagramazero** Revista de Ciência da Informação, Rio de Janeiro, v. 5, n. 4, 2004.

GUINCHAT, C.; MENOU, M. **Introdução geral às ciências e técnicas da informação e documentação.** 2. ed. rev. aum. Brasília: MCT/CNPq/IBICT, 1994.

KATO, Mary A.. **No mundo da escrita: uma perspectiva psicolinguística.** 5. ed. São Paulo: Ática, 1995. (Fundamentos).

LANCASTER, F.W. **Indexação e resumos: teoria e prática.** 2. ed. Brasília: Briquet de Lemos, 2004.

LANGRIDGE, Derek Wilton. **Classificação: abordagem para estudantes de biblioteconomia.** Rio de Janeiro: Interciência, 1977. 126 p. Tradução de Rosali P. Fernandes.

REIS, Daniela Majorie dos. Estudo do contexto de bibliotecas universitárias pela abordagem da cognição profissional de catalogadores e referente à estrutura do documento livro. 2009. 116f. Trabalho de Conclusão de Curso de Biblioteconomia (bacharelado em biblioteconomia) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, 2009. Com bolsa de iniciação científica CNPq/PIBIC de janeiro de 2007 a janeiro de 2010.

SILVEIRA, Telma Jaqueline Dias. **Leitura documentária para catalogação de assunto em ambiente de biblioteca universitária: análise de um modelo de leitura.** 2006. 86f. Trabalho de Conclusão de Curso (Graduação em Biblioteconomia) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2006.

TAYLOR, Arlene G. **The organization of the information.** 2. ed. Westport: Libraries Unlimited, 2004. 417 p.

MODELAGEM E REPRESENTAÇÃO SEMÂNTICA DE DADOS GOVERNAMENTAIS ABERTOS DA PREVIDÊNCIA SOCIAL BRASILEIRA

MODELING AND SEMANTIC REPRESENTATION OF OPEN GOVERNMENT DATA OF BRAZILIAN SOCIAL SECURITY

Durval Vieira Pereira
Carlos Henrique Marcondes

Resumo: Elabora um modelo conceitual dos dados sobre acidentes do trabalho para publicação dos dados governamentais mantidos pela Previdência Social. Propõe um vocabulário sobre acidentes do trabalho e codifica em RDF uma amostra dos dados publicados. O modelo é desenvolvido por meio de análise da definição e identificação dos termos de acordo com as categorias das ontologias UFO-B e da DUL e do modelo de Entidade-Relacionamento. Elabora um vocabulário com escolha dos conceitos a serem utilizados; busca por vocabulários para reuso, cria URIs e nome (*namespace*) para os termos, divide o vocabulário em classes (entidades), propriedades (relacionamentos) e elementos (atributos). O modelo é codificado implementando as triplas RDF em um modelo gráfico RDF/XML. Constata a necessidade da elaboração de um vocabulário para descrever os conceitos sobre acidentes do trabalho presentes no modelo conceitual proposto. Representa uma amostra dos dados em RDF, utilizando vocabulário apresentado. Conclui que a elaboração do modelo conceitual aliada a vocabulários específicos e a descrição em RDF pareceu adequada para organizar e fornecer um nível mínimo de semântica aos dados sobre acidente de trabalho da Previdência Social brasileira.

Palavras-chave: Modelagem conceitual. Modelo Entidade-Relacionamento. Dados governamentais abertos. Acidente do trabalho. Previdência Social.

Abstract: This paper will explore a project about the conceptualization model on occupational accidents at Social Security for open government data. The model is developed through analysis of definition and the identification of terms of terms according to the categories of ontologies UFO-B and DUL and Entity-Relationship Model. It develops of a vocabulary and it choices of concepts to be used, it searches for vocabularies for reuse, it creates URIs and a fictitious name (*namespace*) and it divides the vocabulary in classes (entities), properties (relationships) and elements (attributes). The model is encoded implementing the RDF triples in a graphical RDF/XML model. As a result it was observed the need to develop a vocabulary to describe the concepts of occupational accidents present the proposed conceptual model. And it represents a sample of the data in RDF using vocabulary presented. The research concluded that development of the conceptual model together with the specific vocabularies and description in RDF proved the most suitable for organize and provide a minimum level of semantic data on occupational accidents of Brazilian Social Security.

Keywords: Conceptual modeling. Entity-relationship Model. Open Government Data. Occupational accident. Social Security.

1 INTRODUÇÃO

O W3C (2009, p. 39) define dados governamentais abertos (DGA) como sendo a publicação de informações do setor público “em formato bruto aberto, de maneira a torná-los

acessíveis a todos e permitir sua reutilização, como a criação de *mashups*⁸ de dados”. Em outras palavras busca-se a publicação na *Web* de informações do setor público em formato bruto e aberto, legíveis não somente por pessoas, mas também inteligíveis por máquinas, de modo a permitir sua reutilização em novas aplicações digitais desenvolvidas pela sociedade.

Organizar os dados governamentais e facilitar sua recuperação e sua reutilização é um dos objetivos da abertura desses dados, que vai ao encontro de uma das visões sobre o objeto da Ciência da Informação que “é ajudar as pessoas (ou mais amplamente: atores) que estão confusas com a situação do uso do conhecimento (e que ficarão mais confusas por causa do modelo de sociedade pós-moderna)” (WERSIG, 1993, p. 230, tradução nossa).

Acredita-se que publicação dos dados governamentais abertos pode beneficiar as organizações do setor público ou não (por exemplo: empresas privadas, estatísticos, Organizações Não Governamentais - ONGs - e *Hackers*). Isto é possível através da utilização das tecnologias da Web Semântica para a interligação dos dados de outros órgãos para prover valor agregado a seus próprios dados, melhorando assim sua usabilidade e visibilidade.

É necessário possibilitar o reuso, ou seja, desagregar, reprocessar e reaplicar esses dados, permitindo novas interpretações e aplicações. Pensando nisso, o Governo Brasileiro desenvolveu o Portal Brasileiro de Dados Abertos⁹ para reunir a publicação de dados governamentais e possibilitar sua recuperação e sua reutilização por cidadãos interessados.

Nota-se que a maioria das informações disponibilizadas hoje pelo Governo brasileiro, através do Portal Brasileiro de Dados Abertos, está publicada em formatos proprietários ou com tecnologia que restringe seu acesso. Verificou-se que em 4 de março de 2014 o referido portal apresentava apenas 112 (cento e doze) conjuntos de dados publicados, um número pequeno para o número de órgãos públicos e as informações que geram. Ademais, os dados estavam publicados nos seguintes formatos: xls; pdf; xml; csv; zip+cvs; html; zip+txt; zip+sas; zip+xls e json. Dentre estes, os formatos pdf. e zip, por exemplo, não são considerados formatos adequados para publicação de dados abertos, por não serem formatos abertos. Nenhum conjunto de dados é apresentado em formato *Resource Description Framework* (RDF), “formato recomendado pela proposta de dados abertos interligados” (BENNETT; HARVEY, 2009; BERNERS-LEE, 2009). Contudo, sabe-se que o Ministério do

⁸ *Mashup* é definido pelo W3C (2009, p. 39) como sendo “a mistura de dados de dois ou mais aplicativos ou fontes de dados diferentes, produzindo pontos de vista comparativos das informações combinadas”.

⁹ Disponível em: <<http://dados.gov.br/>>. Acesso em: 19 abr. 2014.

Planejamento¹⁰ e o Tribunal de Contas do Estado de São Paulo¹¹, publicaram conjuntos de dados em formato RDF. Porém, estes não são mostrados junto aos outros formatos existentes e nem facilmente recuperados no portal dados.gov.br.

Segundo Berners-Lee (2009), o padrão RDF conecta melhor que qualquer outro modelo, pois utiliza URIs e assim permite ligação de coisas e conceitos; possibilita que os sistemas independentes possam ser interligados; ocasiona a interoperabilidade; e admite a representação dos dados utilizando diferentes vocabulários.

Na visão de Berners-Lee, Hendler e Lassila (2001) o RDF é fundamental para o estabelecimento do primeiro nível de “semântica” inteligível por programas, pois embora a *eXtensible Markup Language* (XML) permita aos usuários adicionar estrutura arbitrária aos seus documentos, não diz nada sobre o que as estruturas significam. Significados são expressos pelo RDF, que codifica em conjuntos de triplas, sendo cada tripla composta por um sujeito, verbo e objeto de uma frase elementar.

As tecnologias da Web Semântica proporcionam o compartilhamento e reutilização de dados de diferentes aplicações, oferecendo tecnologias que descrevem, modelam e permitam consultas a esses dados. Entretanto, faz necessária a curadoria desses dados para fins de recuperação e reuso e, neste caso, são os modelos que dão estrutura e significado aos dados. Esses fatores evidenciam a necessidade de modelos conceituais e conseqüentemente da modelagem das relações conceituais (SALES; SAYÃO; MOTTA, 2012).

Logo, a publicação de dados governamentais abertos deve se basear nas tecnologias da Web Semântica, pois assim os dados receberão significados definidos, com metadados em formato padronizado, possibilitando a compreensão por pessoas e por computadores.

O presente artigo é resultado de pesquisa de Mestrado defendida junto ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal Fluminense. E seu recorte recai sobre a publicação de dados estatísticos sobre acidentes de trabalho, do Ministério da Previdência Social, geridos pela Empresa de Tecnologias e Informação da Previdência Social (Dataprev). Acidente do trabalho é acidente que “ocorre pelo exercício do trabalho a serviço da empresa ou pelo exercício do trabalho aos segurados empregados, trabalhadores avulsos e segurados especiais” (HORVATH JÚNIOR, 2009, p. 6).

O recorte é justificado por esses dados fazem parte da primeira experiência com a publicação de dados abertos realizada pela Empresa de Tecnologias e Informação da

¹⁰ Disponível em: <<http://www.siop.planejamento.gov.br>>. Acesso em: 5 mar. 2014.

¹¹ Disponível em: <<http://portaldocidadao.tce.sp.gov.br>>. Acesso em: 5 mar. 2014.

Previdência Social (Dataprev), guardiã e gestora dos dados da Previdência Social brasileira (DATAPREV, 2012).

Para isto foi desenvolvido um modelo conceitual sobre acidente de trabalho, que serve de base para a utilização das tecnologias da Web Semântica para a publicação dos dados governamentais abertos da Previdência Social. A partir do modelo é proposto um vocabulário sobre acidentes do trabalho e ainda codificar em RDF uma amostra dos dados publicados.

2 METODOLOGIA

A metodologia foi dividida em três etapas:

➤ **Elaboração de um modelo conceitual sobre acidente do trabalho**

- a. Análise da definição e identificação dos termos com base na Teoria do Conceito, desenvolvida por Dahlberg (1978). Dahlberg (1981; 1983) afirma que as definições são “pressupostos indispensáveis na argumentação e nas comunicações verbais e que constituem elementos necessários na construção de sistemas científicos” (DAHLBERG, 1978, p. 106). Dahlberg (1995) reconhece no conceito uma unidade de conhecimento, o conceitua de forma objetiva, como uma manifestação linguística desta unidade de conhecimento, o termo, que se refere a um único objeto do mundo e que por sua vez sintetiza seu significado. É o assim chamado princípio da *monoreferencialidade*. Em termos metodológicos, as definições de acidente de trabalho foram analisadas para determinar a que entidades e relações os termos nelas encontrados se referiam. Ao desconfiar que mais de uma palavra da definição pudesse formar um único conceito, verificou-se se cada uma dessas palavras separadamente teria um significado diferente do significado pretendido quando elas eram coordenadas formando um único termo.
- b. Classificação dos termos identificados de acordo com as categorias das ontologias *Unified Foundational Ontology for perdurants* (UFO-B) e da DOLCE¹²+DnS *Ultralite* (DUL), verificando semelhanças e diferenças, contribuindo para a identificação dos elementos (objeto ou relação) do modelo a ser construído;
- c. Classificação dos termos de acordo com o modelo de Entidade-Relacionamento, desenvolvido por Chen (1976).
- d. Elaboração de um diagrama de Entidade-Relacionamento para a definição. A escolha do Modelo Entidade-Relacionamento é justificada pela sua proximidade com o RDF,

¹² Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE) e Descriptions and situations (DnS).

padrão indicado para descrição de dados na Web Semântica. Chen (2002) ao falar dos possíveis usos e relações do modelo Entidade-Relacionamento o associa ao RDF, afirmando que existem algumas semelhanças e diferenças entre RDF e do modelo ER. Entretanto, o RDF, já é considerado um membro da família da modelagem Entidade-Relacionamento, em que os dados estruturados como gráficos rotulados podem ser trocados através de documentos XML (W3C, 1999).

- e. Apresentação um modelo conceitual abrangendo a área de acidente do trabalho no domínio da Previdência Social.

➤ **Elaboração de um vocabulário sobre acidente do trabalho**

- a. Escolha dos conceitos do modelo conceitual a serem utilizados na elaboração do vocabulário. Não serão contemplados todos os conceitos neste momento, apenas aqueles relacionados diretamente aos dados já publicados pela Previdência Social;
- b. Busca por vocabulários já existentes que possam contemplar essas entidades, relacionamentos e atributos;
- c. Criação de URIs fictícios para os conceitos a serem criados;
- d. Definição de um nome (*namespace*) para o vocabulário;
- e. Descrição dos conceitos que farão parte do vocabulário proposto;
- f. Divisão do vocabulário em classes (entidades), propriedade (relacionamentos) e elementos (atributos).

➤ **Codificação em RDF de uma amostra dos dados publicados pela Previdência Social sobre acidente do trabalho**

- a) Definição das triplas RDF com base no modelo conceitual proposto.
- b) Representação das triplas RDF em quadros, baseados em na criação de tabelas para descrição de banco de dados lógicos (XU; LEE; KIM, 2010) para descrever as entidades e relações e convertê-las para triplas RDF;
- c) Substituição dos elementos do quadro por suas respectivas URI, de acordo com o vocabulário proposto;
- d) Implementação das triplas RDF nos dados selecionados em um modelo gráfico RDF/XML.

3 RESULTADOS

Os resultados encontrados foram divididos em três fases. A primeira referente à concepção do modelo conceitual. A segunda à elaboração de uma proposta de vocabulário

sobre acidente de trabalho, a ser utilizado na terceira fase, que é a descrição em RDF dos dados sobre acidentes do trabalho.

3.1 Modelo conceitual sobre acidente de trabalho associado aos dados da previdência social

Com base nas definições analisadas sobre acidente do trabalho encontradas em Brasil (1991), Anuário Estatístico da Previdência Social (2012), *Organización Internacional del Trabajo* (2011) e Horvath Júnior (2009) se chegou a uma definição única a ser utilizada para a elaboração do modelo conceitual.

Acidente do trabalho: acidente sofrido pelos segurados empregados, trabalhadores avulsos e segurados especiais no exercício do trabalho a serviço da empresa, que cause lesão corporal ou perturbação funcional, capaz de gerar morte, a perda ou a redução, permanente ou temporária, da capacidade laborativa, regulamentado por art. 19 da Lei n. 8.213/91. Classificados em acidente típico, acidente de trajeto e acidente atípico, que cause doença profissional ou doença do trabalho.

A definição acima foi analisada utilizando a Teoria do Conceito de Dahlberg (1978) para análise das definições e identificação dos termos, estes sendo palavra ou expressão que tenha sempre um único referente na realidade, ou seja, o princípio da monoreferencialidade; os termos identificados desta maneira nas definições estão marcados com “<” e “>”. Também classificou-se os termos identificados de acordo com as categorias das ontologias *Unified Foundational Ontology for perdurants* (UFO-B) e da DOLCE+DnS *Ultralite* (DUL). E a partir disso, classificou-se os termos de acordo com o modelo de Entidade-Relacionamento, desenvolvido por Chen (1976), a fim de elaborar um diagrama Entidade-Relacionamento para o domínio estudado.

QUADRO 1 - Análise da definição sobre acidente do trabalho.

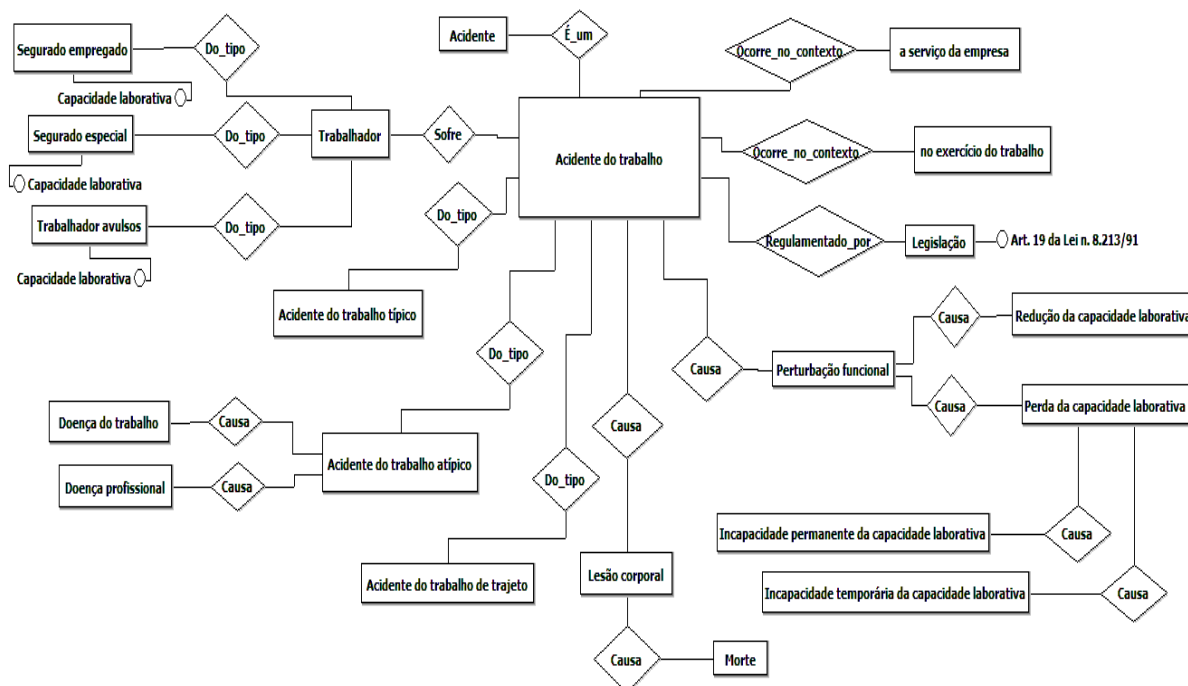
Definição	<Acidente do trabalho>: <acidente> <sofrido> pelos <segurados empregados>, <trabalhadores avulsos> e <segurados especiais> <no exercício do trabalho> <a serviço da empresa>, que <cause> <lesão corporal> ou <perturbação funcional>, capaz de <gerar> <morte>, <perda> ou a <redução>, <permanente> ou <temporária>, da <capacidade laborativa>, <regulamentado> por <art. 19 da Lei n. 8.213/91>. <Classificados> em <acidente típico>, <acidente de trajeto> e <acidente atípico>, que <cause> <doença profissional> ou <doença do trabalho>.			
Classificação da definição (Dahlberg)	Definição genérica			
Conceitos retirados da definição	Categoria da UFO-B	Categoria da DUL	Categoria E-R	Observação

Acidente do trabalho	Evento	Evento	E	
acidente	Evento	Evento	E	Categoria superior a acidente do trabalho
segurados empregados	Objeto	Objeto / Agente	E	O objeto é quem sofre o acidente do trabalho
trabalhadores avulsos	Objeto	Objeto / Agente	E	O objeto é quem sofre o acidente do trabalho
segurados especiais	Objeto	Objeto / Agente	E	O objeto é quem sofre o acidente do trabalho
pelo exercício do trabalho	Situação	Situação	E	Relaciona o participante ao evento
a serviço da empresa	Situação	Situação	E	Relaciona o participante ao evento
lesão corporal	Evento	Evento	E	Relação entre eventos
perturbação funcional	Evento	Evento	E	Relação entre eventos
morte	Evento	Evento	E	Relação causal entre eventos: “lesão corporal” e “morte”
redução da capacidade para o trabalho	Evento	Evento	E	Relação causal entre eventos: “Perturbação funcional” e “redução da capacidade para o trabalho”
perda da capacidade para o trabalho	Evento	Evento	E	Relação causal entre eventos: “Perturbação funcional” e “perda da capacidade para o trabalho”. Que possui subcategorias: “Incapacidade permanente da capacidade para o trabalho” ou “Incapacidade temporária da capacidade para o trabalho”
art. 19 da Lei n. 8.213/91	Objeto	Objeto / Objeto social	E	Documento regulamenta o acidente do trabalho
capacidade laborativa			A	Atributo do trabalhador
acidente típico	Evento	Evento	E	Tipos de acidentes do trabalho
acidente atípico	Evento	Evento	E	Tipos de acidentes do trabalho
acidente de trajeto	Evento	Evento	E	Tipos de acidentes do trabalho
doença profissional	Evento	Evento	E	Sinônimo: ergopatia
doença do trabalho	Evento	Evento	E	Sinônimo: mesopatia
sofrido			R	Relação associativa
Cause			R	Relação causal
gerar			R	Relação causal
regulamentado			R	Relação associativa
classificados			R	Relação gênero/espécie

Fonte: PEREIRA (2014).

A partir do quadro de análise, elaborou-se um diagrama ER para ilustrar o modelo conceitual proposto.

FIGURA 1 - Diagrama ER da definição sobre acidente do trabalho.



Fonte: PEREIRA (2014).

Mesmo com um modelo conceitual sobre acidente do trabalho proposto, sentiu-se falta da inclusão de outros elementos, não apresentados nas definições estudadas. No âmbito na Previdência Social, o acidente do trabalho é utilizado para provimento de benefícios aos trabalhadores acidentados. Por esta razão, buscou-se as entidades complementares para a elaboração do modelo proposto. Desta forma, cabe interligar o modelo proposto sobre acidente do trabalho aos dados já publicados e a outros que se pretende publicar da Previdência Social.

Desta forma, coletaram-se os termos do Anuário estatístico de acidente do trabalho (2012) e Anuário estatístico da Previdência Social (2011), os termos que associassem os dados publicados ao acidente do trabalho. Os termos selecionados estão descritos no quadro abaixo:

QUADRO 2 - Termos sobre acidente do trabalho aliados aos dados da Previdência Social.

Fontes	Anuário estatístico de acidente do trabalho (2013) Anuário estatístico da Previdência Social (2011)			
Conceitos	Categoria da UFO-B	Categoria da DUL	Categoria E-R	Observação
Acidente do trabalho com CAT (Comunicação de Acidente de Trabalho)	Objeto	Objeto / Objeto social	E	Acidente registrado (documento)

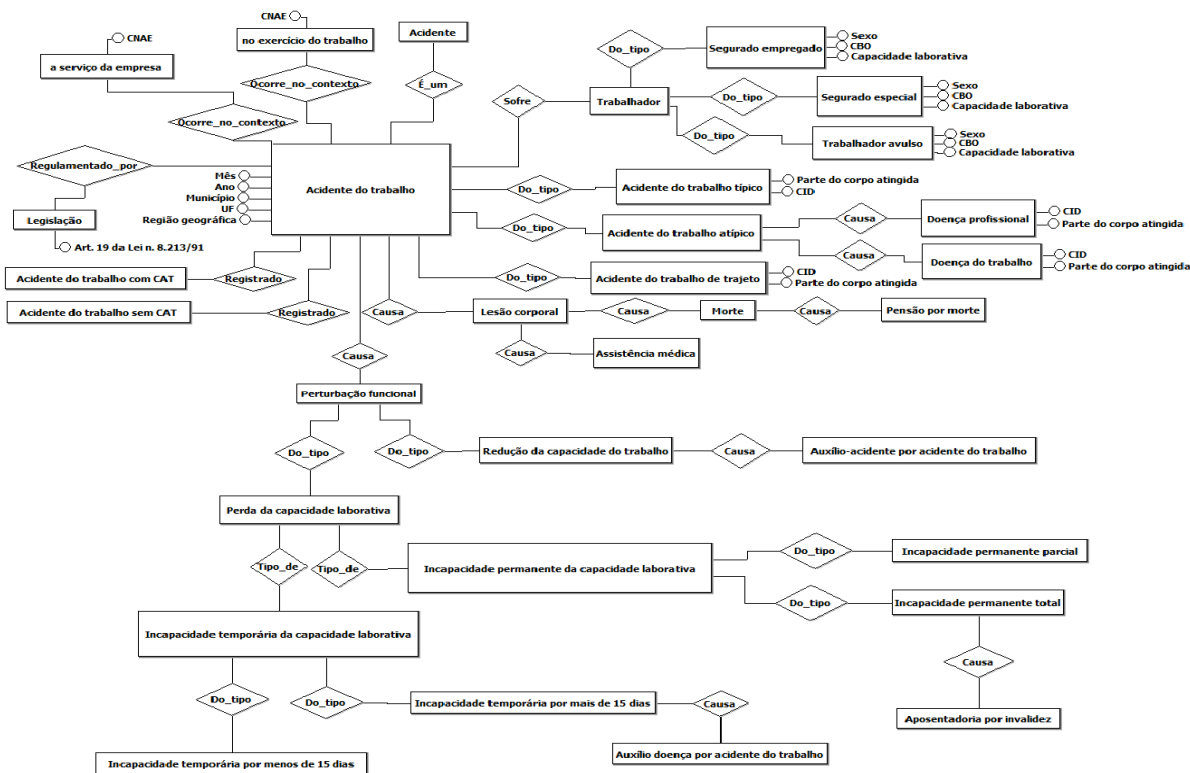
Acidente do trabalho sem CAT (Comunicação de Acidente de Trabalho)	Objeto	Objeto / Objeto social	E	Acidente não registrado no INSS. O acidente é identificado por meio de um dos possíveis nexos: nexo técnico profissional/trabalho, nexo técnico epidemiológico previdenciário ou nexo técnico por doença equiparada a acidente do trabalho. Esta identificação é feita pela nova forma de concessão de benefícios acidentários.
Ano	Intervalo temporal		A	Atributo da entidade “Acidente do trabalho”. Ano que ocorreu o acidente do trabalho
Aposentadoria por invalidez	Objeto	Objeto / Objeto social	E	Benefício concedido ao trabalhador em caso de comprovação de perda permanente da capacidade laborativa
Assistência médica	Objeto	Objeto / Objeto social	E	Benefício concedido ao trabalhador em caso de lesão corporal
Auxílio doença por acidente do trabalho	Objeto	Objeto / Objeto social	E	Benefício concedido ao trabalhador em caso de interrupção por mais de 15 dias das atividades do trabalho em função da perda da capacidade laborativa
Auxílio-acidente por acidente do trabalho	Objeto	Objeto / Objeto social	E	Benefício concedido ao trabalhador em caso de redução da capacidade do trabalho
CBO (classificação Brasileira de Ocupações)			A	Atributo que identifica a ocupação do trabalhador
CID (Classificação Internacional de Doenças)			A	Atributo dos acidentes, que identifica a doença/lesão causada
CNAE (Classificação Nacional de Atividades Econômicas)			A	Atributo que identifica a atividade econômica do empregador
Incapacidade permanente parcial	Evento	Evento	E	Sequela definitiva que implique em redução da capacidade laboral do trabalhador
Incapacidade permanente total	Evento	Evento	E	Incapacidade permanente e total para o exercício de qualquer atividade laborativa
Incapacidade temporária por mais de 15 dias	Evento	Evento	E	Interrupção por mais de 15 dias das atividades do trabalho em função da perda da capacidade laborativa
Incapacidade temporária por menos de 15 dias	Evento	Evento	E	Interrupção por até 15 dias das atividades do trabalho em função da perda da capacidade laborativa
Mês	Intervalo temporal		A	Atributo da entidade “Acidente do trabalho”. Mês que ocorreu o acidente do

				trabalho	
Município		Lugar	A	Atributo da entidade “Acidente do trabalho”. Cidade em que ocorreu o acidente do trabalho	
Parte do corpo atingida			A	Atributo dos acidentes, que identifica a parte do corpo do trabalhador lesionado em função do acidente	
Pensão por morte	por	Objeto	Objeto / Objeto social	E	Benefício concedido ao dependente do trabalhador em caso de sua morte
Região geográfica			Lugar	A	Atributo da entidade “Acidente do trabalho”. Região geográfica brasileira na qual ocorreu o acidente do trabalho
Sexo				A	Atributo que identifica o sexo do trabalhador
UF (Unidade Federativa)			Lugar	A	Atributo da entidade “Acidente do trabalho”. Estado brasileiro no qual ocorreu o acidente do trabalho

Fonte: Pereira (2014).

Na FIG. 2, pode ser observado o mesmo modelo (FIG. 1), acrescido de atributos e de outras entidades e relações presentes nos dados da Previdência Social, publicados pela Dataprev, e também presentes na seção IV do Anuário estatístico da Previdência Social, sobre acidente do trabalho.

FIGURA 2 - Diagrama ER sobre acidente do trabalho aliados aos dados da Previdência Social



Fonte: Pereira (2014).

O modelo conceitual proposto visa contribuir para melhoria semântica dos dados abertos publicados pelo Ministério da Previdência Social. Contribuindo ainda para que outros estudos busquem modelos para seus respectivos domínios dentro do Governo Brasileiro. Aliar o modelo conceitual a elaboração de um vocabulário coopera para a fundamentação semântica dos dados governamentais abertos publicados sobre acidente de trabalho. Pois, os “vocabulários baseados em linguagens de domínio, disciplina ou comunidade, ajudam a resolver problemas relacionados à sobrecarga de informações digitais, bem como auxilia em sua recuperação” (MÉNDEZ; GREENBERG, 2012, p. 237).

3.2 ELABORAÇÃO E REUSO DE VOCABULÁRIOS: A QUESTÃO SOBRE ACIDENTE DO TRABALHO

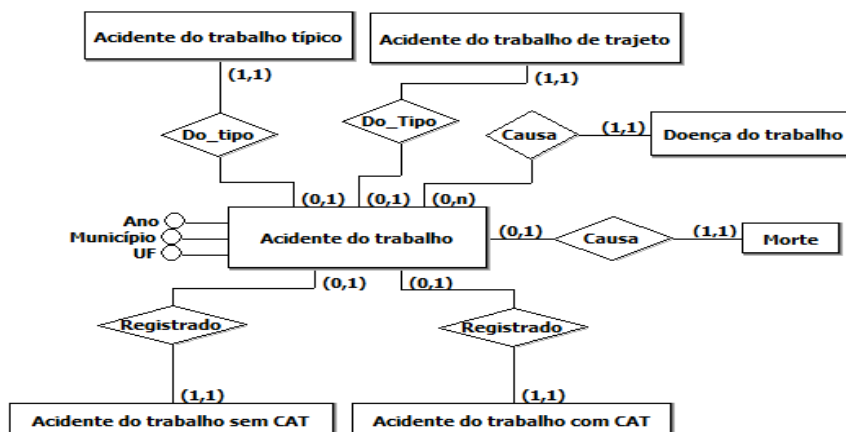
Buscou-se construir um vocabulário a ser usado para dar significado às triplas RFD para publicação de dados governamentais abertos, utilizando os moldes dos vocabulários encontrados no portal do *Linked Open Vocabularies* (LOV)¹³. Elaborado para ser “um compilador de ontologias (vocabulários). Seu objetivo é fornecer acesso a vocabulários, tornando explícitas as representações entre as entidades (objetos)” (VANDENBUSSCHE, [2014?]).

Para aplicação do modelo proposto, restringiu-se a amostragem aos dados da Seção I subseção D do Anuário Estatístico de Acidente do Trabalho, conteúdo que trata de Comunicação de Acidente de Trabalho (CAT), suas consequências, por localização geográfica das ocorrências (município) e seus respectivos quantitativos.

Assim, foram descritos os seguintes dados: ano de ocorrência do acidente do trabalho; dados sobre o município, onde o acidente do trabalho ocorreu (código da cidade, fornecido pelo IBGE, nome do município e Unidade da Federação, a qual pertence o município); e quantidades de acidentes, divididos por tipo, causas e registro (acidente do trabalho do tipo típico, acidente do trabalho do tipo trajeto, acidente do trabalho causou doença do trabalho, acidente do trabalho causou morte e acidente do trabalho registrado sem CAT).

Como o modelo proposto pretende atender a um conjunto de dados maiores, ele apresenta classes e subclasses, além de relações não necessárias, neste momento, para descrição dos dados a serem descritos em RDF. Por essa razão, decidiu-se recortar o referido modelo, sintetizando sua representação. Abaixo se pode observar o diagrama Entidade-Relacionamento apenas dos dados já publicados pela Dataprev.

¹³ Disponível em: <<http://lov.okfn.org/dataset/lov/>>. Acesso em: 6 jul. 2014.

FIGURA 3 - Proposta de diagrama ER dos dados publicados pela Dataprev¹⁴.

Fonte: PEREIRA (2014).

Estes dados são disponibilizados nos formatos XML, JSON ou CSV. Decidiu-se trabalhar com o formato XML, por ser o recomendado pela W3C para ser utilizado junto ao RDF no aperfeiçoamento de tecnologias da Web Semântica.

FIGURA 4 - XML de acidentes do trabalho ocorridos na cidade do Rio de Janeiro

```
<acidentes_de_trabalho>
<registro>
<municipio cod_ibge="330000" uf="RJ">RIO DE JANEIRO</municipio>
<quantidade>
<sem_cat>9</sem_cat>
<com_cat>
<tipicos>17</tipicos>
<trajeto>17</trajeto>
<doenca>0</doenca>
</com_cat>
<obitos>0</obitos>
</quantidade>
</registro>
</acidentes_de_trabalho>
```

Fonte: DATAPREV (2012).

Na FIG. 4, foram selecionados os dados referentes a acidente do trabalho ocorridas no ano de 2009, último ano disponibilizado. Dentre os dados estatísticos de todas as cidades, optou-se pelos dados de apenas uma cidade, já que o processo seria o mesmo para todas. A cidade escolhida foi o Rio de Janeiro.

Ao selecionar os dados a serem representados em RDF, verificou-se que seria necessária a elaboração de um vocabulário específico para descrever os conceitos sobre acidentes do trabalho como forma de enriquecer a representação dos dados da Previdência em

¹⁴ No diagrama proposto por Pereira (2014) não foi utilizado a cardinalidade, porém para fim de aplicação, no diagrama da figura 2 foram inseridos a cardinalidade.

RDF. Fato este justificado, por não encontrarmos instrumentos de representação, específicos sobre acidentes do trabalho, capazes de serem utilizados para o fim proposto.

A partir das classes definidas, realizou-se uma pesquisa no LOV, com o objetivo de verificar se haveria alguma classe ou relação já existente em ontologias (vocabulários), que representassem as classes ou relações sobre acidente do trabalho. Tal ação visava o reuso de alguns vocabulários, já existentes e consolidados, para a interligação dos dados, atendendo um dos princípios do *Linked Data*. Além de fornecer um formalismo semântico maior a proposta aqui apresentada. Entretanto, o resultado foi negativo, não encontrando nenhuma classe capaz de atender aos significados no domínio aqui estudado.

Antevendo, como todo vocabulário precisaria possuir um URI para suas classes, além da necessidade de descrição destas classes em RDF, foi criado um URI fictício, utilizado apenas para exemplificação neste trabalho. Para sua elaboração foi gerado um recurso eletrônico (“<http://example.org/>”) e após a barra deste endereço a classe a ser descrita. Foi dado o prefixo “acitra” para representar o *namespace* do Vocabulário de Acidente do Trabalho, também em caráter de experimentação.

QUADRO 3 - Classe do objeto

Classe	acitra:acidenteTrabalho
Namespace	Acitra
URI	http://example.org/acidenteTrabalho
Label ¹⁵	acidenteTrabalho
Definição	Acidente que ocorre pelo exercício do trabalho a serviço da empresa ou pelo exercício do trabalho dos segurados especiais, provocando lesão corporal ou perturbação
Classe	acitra:acidenteTrabalhoTípico
Namespace	Acitra
URI	http://example.org/acidenteTrabalhoTípico
Label	acidenteTrabalhoTípico
Definição	Acidente decorrente da característica da atividade profissional de risco desempenhada pelo acidentado
Classe	acitra:acidenteTrabalhoTrajeto
Namespace	Acitra
URI	http://example.org/acidenteTrabalhoTrajeto
Label	acidenteTrabalhoTrajeto
Definição	Acidente que cause a morte ou produza lesões corporais e ocorra no

¹⁵ No item Label foram desconsiderados os artigos e preposições. Permaneceram apenas as preposições nos Label “acidenteTrabalhoSemCAT” e “acidenteTrabalhoComCAT” para diferenciá-los semanticamente.

	percurso direto entre o lugar de trabalho e a residência, o lugar no que o trabalhador geralmente almoça ou recebe sua remuneração
Classe	acitra:doençaTrabalho
Namespace	Acitra
URI	http://example.org/doençaTrabalho
Label	doençaTrabalho
Definição	Doença relacionada diretamente ao ambiente laboral, ou seja, aos riscos ambientais ao que o trabalhador é exposto
Classe	acitra:morte
Namespace	Acitra
URI	http://example.org/morte
Label	Morte
Definição	Corresponde a quantidade de segurados que faleceram em função do acidente do trabalho.
Classe	acitra:acidenteTrabalhoSemCAT
Namespace	Acitra
URI	http://example.org/acidenteTrabalhoSemCAT
Label	acidenteTrabalhoSemCAT
Definição	Acidente cuja Comunicação de Acidentes do Trabalho (CAT) não foi cadastrada no INSS.
Classe	acitra:acidenteTrabalhoComCAT
Namespace	Acitra
URI	http://example.org/acidenteTrabalhoComCAT
Label	acidenteTrabalhoComCAT
Definição	Acidente cuja Comunicação de Acidentes do Trabalho (CAT) foi cadastrada no INSS.

Fonte: Pereira (2014).

Um segundo passo, foi buscar, também no LOV, vocabulários capazes de representar as relações mostradas no diagrama ER. Pesquisou-se por propriedades do objeto (*owl:ObjectProperty*) com significados iguais ao que se propunha no modelo.

Foram encontrados resultados satisfatórios, que evitou a criação ou utilização de propriedades do objeto ficcionais. Assim, os termos utilizados para representarem as relações foram substituídos por termos de vocabulários já existentes. O termo “Do_Tipo” foi substituído pelo termo “*Type*”, do vocabulário Dublin Core. Já o termo “Causa” foi substituído pelo termo “*Cause*”, Schema.Org. E, o termo “*hasDocumentation*”, do vocabulário Poder Vocabulary substituiu o termo “Registrado”. Desta forma, abaixo são descritas as propriedades do objeto.

QUADRO 4 - Propriedades do objeto

Propriedade	dc:type
Namespace	dc/elements
URI	http://purl.org/dc/elements/1.1/type
Label	Type
Definição	Termos que descrevem categorias gerais, funções, gêneros ou níveis de agregação de conteúdo ¹⁶ .
Propriedade	dc:cause
Namespace	schema.org
URI	http://schema.org/cause
Label	Cause
Definição	Uma causa subjacente. Mais especificamente, um dos agente(s) é responsável pelo processo fisiopatológico que, eventualmente, resulta na ocorrência ¹⁷ .
Propriedade	poder:hasDocumentation
Namespace	Poder
URI	http://dev.poderopedia.com/vocb/hasDocumentation
Label	hasDocumentation
Definição	Associa um agente com a documentação relacionada ¹⁸ .

Fonte: Pereira (2014).

No caso apresentado, a única classe a apresentar atributos é a classe “Acidente do trabalho”. Pesquisou-se no LOV os “*Objectdata*”, que correspondessem aos atributos selecionados nesta pesquisa. Assim, também foram encontrados vocabulários possíveis de serem reutilizados, não sendo necessária a sua criação.

O atributo “ano” foi substituído pelo atributo “Date”, do vocabulário Dublin Core. O “Município” foi substituído por “City”, do vocabulário Place, e deste mesmo vocabulário, “State” substituiu “UF”. A seguir, pode se verificar estes atributos e sua descrição.

QUADRO 5 - Atributos

Elemento	dc:date
Namespace	dc/elements
URI	http://purl.org/dc/elements/1.1/date
Label	Date
Definição	Representa o ano em que ocorreu o acidente do trabalho.
Domínio	acitra:acidenteTrabalho

¹⁶ Pertence ao vocabulário **DC – Dublin Core Metadata Set**. Seus elementos são amplos e genéricos, utilizáveis para descrever uma ampla gama de recursos (DUBLIN CORE METADATA INITIATIVE, 2012).

¹⁷ Proveniente do vocabulário **Schema.org**, utilizado nos motores de busca de alguns motores de busca na Web, que dependem de marcação schema.org para melhorar a exibição de resultados (SCHEMA.ORG, [2014?]).

¹⁸ Retirado do **Poder Vocabulary**, utilizado para expressar informações acerca de pessoas e suas relações com organizações e empresas (PODER, [2014?]).

Range	<i>xsd</i> ¹⁹ : <i>datetime</i>
Elemento	places:City
Namespace	Places
URI	http://purl.org/ontology/places#City
Label	City
Definição	Representa o município (cidade brasileira) na qual ocorreu o acidente do trabalho.
Domínio	acitra:acidenteTrabalho
Range	<i>name:nome</i> ²⁰ .
Elemento	places:State
Namespace	Places
URI	http://purl.org/ontology/places#State
Label	State
Definição	Representa a Unidade Federativa (Estado brasileiro) na qual ocorreu o acidente do trabalho.
Domínio	acitra:acidenteTrabalho
Range	<i>name:uf</i>

Fonte: Pereira (2014).

Assim sendo, consegue-se mostrar, com os exemplos acima, que é possível se chegar a um vocabulário partindo de um modelo conceitual. Espera-se, no futuro, transformar todo o modelo conceitual proposto em um vocabulário consistente e formalizado de forma a conseguir sua publicação, visando sua reutilização.

3.3 DESCRIÇÃO DOS DADOS SOBRE ACIDENTE DO TRABALHO EM RDF

Com a definição das classes, das relações (propriedades do objeto) e dos atributos (propriedade de dados), buscou-se a partir do diagrama Entidade-Relacionamento estabelecer *statements*, que são “uma espécie de declaração de um recurso contendo um nome, uma propriedade e um valor agregado a ela” (NOLETO; BRITO, 2003, p. 114), formando a tripla RDF (sujeito, predicado e objeto), conseguindo representar a interligação entre o recurso, suas propriedades e seus valores.

Apesar de se perceber algumas semelhanças entre o modelo ER e o RDF, a relação entre eles não é fácil, quando se pretende passar de um modelo para o outro. Chen (2002) e Berners-Lee (1998) apontam semelhanças e diferenças, porém não apresentam uma metodologia para que os dois modelos possam trabalhar de forma cooperativa.

¹⁹ XSD (*XML Schema Definition*) é um arquivo que contém definições na linguagem XML Schema, utilizado para descrever a estrutura de um documento XML (SPERBERG-MCQUEEN; THOMPSON, 2010).

²⁰ Retirado de *Places Ontology*, usado para descrever os locais de interesse geográfico (PLACES..., [2014?]).

Buscou-se na literatura da área da Ciência da Computação experiências na associação do modelo ER e RDF para descrição de dados. Os pesquisadores Xu, Lee e Kim (2010) utilizam o diagrama ER para armazenar dados RDF em tabelas relacionais distintas. Os autores acreditam que com essa abordagem facilitaria a pesquisa e a atualização de quaisquer recursos em bases relacionais.

Esses autores buscaram descrever RDFS em um diagrama ER, pois o diagrama apresenta relações de dados RDF usando dados de próprio esquema. Para tal, foram convertidas todas as classes de entidades e as propriedades das relações para RDF/RDFS. Isto foi possível, pois criaram tabelas para os esquemas: uma tabela para armazenar as entidades (tabela de Classes) e os relacionamentos (tabela de propriedades do objeto).

Rambayon (2012) acredita que a mesma metodologia utilizada por Xu, Lee e Kim (2010), pode ser aplicada para ter resultados inversos, ou seja, a partir de um diagrama ER se chegar a descrições em RDF.

Vale lembrar que, “o objetivo da modelagem conceitual é obter uma descrição abstrata, independente de implementação em computador, dos dados que serão armazenados no banco de dados” (HEUSSER, 1998, p. 11). Desta maneira, deve-se deixar claro que este trabalho está na fase de abstração, ou seja, certos detalhes são deliberadamente omitidos do modelo. E a escolha dos detalhes/características para omitir é “feita considerando-se tanto a aplicação pretendida da abstração e também seus usuários” (SMITH; SMITH, 1977, p. 105). No caso específico deste estudo, a abstração também se deve ao fato da representação dos dados sobre acidentes do trabalho ser uma experiência, na qual se verificará uma possibilidade de modelagem e aplicação do modelo nos dados.

Buscou-se nas entidades e relacionamentos estudados no decorrer do trabalho, para a criação das sentenças ou triplas RDF, conforme abaixo.

QUADRO 6 - Triplas de RDF propostas com base no diagrama ER

Sujeito	Predicado	Objeto
Acidente do trabalho	Do_tipo	Acidente do trabalho típico
Acidente do trabalho	Do_tipo	Acidente do trabalho de trajeto
Acidente do trabalho	Causa	Doença do trabalho
Acidente do trabalho	Causa	Morte
Acidente do trabalho	Registrado	Acidente do trabalho sem CAT
Acidente do trabalho	Registrado	Acidente do trabalho com CAT

Fonte: PEREIRA (2014).

A Web Semântica se diferencia da web tradicional por se preocupar mais com seus endereços, ou seja, tornando seus endereços permanentes, utilizando para tal tarefa os URIs.

Como lembra Noletto e Brito (2003, p. 114) “todos os recursos devem estar nomeados e identificados por um URI”, isto permite a identificação dos recursos. Por esta razão, será repetido o quadro acima, substituindo os termos por seus respectivos URIs (vide vocabulário aqui proposto).

QUADRO 7 - Triplas de RDF propostas utilizando URIs

Sujeito	Predicado	Objeto
http://example.org/acidenteTrabalho	http://purl.org/dc/elements/1.1/type	http://example.org/acidenteTrabalhoTípico
http://example.org/acidenteTrabalho	http://purl.org/dc/elements/1.1/type	http://example.org/acidenteTrabalhoTrajeto
http://example.org/acidenteTrabalho	http://schema.org/cause	http://example.org/doençaTrabalho
http://example.org/acidenteTrabalho	http://schema.org/cause	http://example.org/morte
http://example.org/acidenteTrabalho	http://dev.poderopedia.com/vocab/hasDocumentation	http://example.org/acidenteTrabalhoSemCAT
http://example.org/acidenteTrabalho	http://dev.poderopedia.com/vocab/hasDocumentation	http://example.org/acidenteTrabalhoComCAT

Fonte: PEREIRA (2014).

Apesar da tabela acima mostrar como ficariam as sentenças (*statements*) RDF, para se construir e disponibilizar os dados sobre acidentes do trabalho em RDF, faz-se necessário que estes estejam em forma textual. Seguindo as recomendações do W3C (2009), o RDF será expresso textualmente na linguagem RDF/XML, que possibilita a descrição do RDF em formato XML.

FIGURA 5 - Representação dos dados selecionados em RDF²¹

```

<?xml version="1.0"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns"
  xmlns:acid="http://example.org/"
  xmlns:dc="http://purl.org/dc/elements/1.1"
  xmlns:vocab="http://dev.poderopedia.com/vocab/schema"
  xmlns:schema=" schema.org"
  xmlns:place=" http://purl.org/ontology/places#">
  <RDF:DescriptionRdf:about="http://example.org/acidenteTrabalho">
    <dc:date>2009</dc:date>
    <place:city>330000 "Rio de Janeiro"</place:city>
    <place:state>RJ</place:state>
    <acid:acidenteTrabalhoSemCAT>9</acid:/acidenteTrabalhoSemCAT>
    <acid:acidenteTrabalhoComCAT>34</acid:/acidenteTrabalhoComCAT>
    <acid:acidenteTrabalhoTípico>17</acid:/acidenteTrabalhoTípico>
    <acid:acidenteTrabalhoTrajeto>17</acid:/acidenteTrabalhoTrajeto>
    <acid:doençaTrabalho>0</acid:/doençaTrabalho>
    <acid:morte>0</acid:/morte>
  </rdf:Description>
</rdf:RDF>
<?Xml Version="1.0 "?>

```

Fonte: Pereira (2014).

Conseguindo-se aplicar as triplas RDF geradas à faixa dos dados sobre acidentes do trabalho, publicados pela Dataprev, torna-se mais fácil a representação de todo o conjunto de dados também em RDF. Basta se desenvolver uma tabela das triplas e armazená-las no banco de dados relacional, onde os dados sobre acidente do trabalho estão disponibilizados.

Conclui-se que, se os dados são ofertados nos formatos XML, JSON ou CSV, eles também poderão ser disponibilizados em RDF, obtendo um nível semântico não permitido pelos outros formatos atualmente utilizados.

5 CONSIDERAÇÕES FINAIS

O presente estudo se propôs a oferecer uma pequena contribuição para que o Governo Eletrônico, em destaque o portal dados.gov.br, tenha uma alternativa para publicação de dados governamentais abertos, neste caso dados da Previdência Social.

Espera-se que este trabalho possa contribuir com idéias e práticas para que a publicação desses e de outros dados seja realizada, utilizando princípios dos dados governamentais abertos. De tal forma a possibilitar a criação, pela própria sociedade, de novos serviços e informações que contribuirão para a construção de uma cidadania, no qual os

²¹ RDF validado pelo *W3C RDF validation service*. Disponível em: <<http://www.w3.org/RDF/Validator/>>. Acesso em: 28 abr. 2014.

cidadãos terão maior acesso às informações e, conseqüentemente, a construção de uma sociedade mais justa, participativa e com maiores oportunidades.

Por fim, no contexto apresentado neste trabalho, a elaboração do modelo conceitual e a descrição em RDF pareceu adequado para organizar e fornecer um nível mínimo de semântica aos dados sobre acidente de trabalho da Previdência Social brasileira, com vistas a promover a organização e facilitar a recuperação e reuso desses dados.

REFERÊNCIAS

ANUÁRIO estatístico da previdência social: AEPS 2012. Brasília, v. 20, 2013.

ANUÁRIO estatístico de acidentes do trabalho: AEAT 2010. Brasília, v. 1, 2011.

BERNERS-LEE, T. **Putting government data online**. 2009. Disponível em: <<http://www.w3.org/DesignIssues/GovData.html>>. Acesso em: 11 jun. 2014.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. **The Semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities**. 2001. Disponível em: <http://www-sop.inria.fr/acacia/cours/essi2006/Scientific%20American_%20Feature%20Article_%20The%20Semantic%20Web_%20May%202001.pdf>. Acesso em: 13 mar. 2014.

BERNERS-LEE, T. **Relational databases on the Semantic Web**. 1998. Disponível em: <<http://www.w3.org/DesignIssues/RDB-RDF.html>>. Acesso em: 2 mar. 2014.

BRASIL. Lei nº 8.213, de 24 de julho de 1991. Dispõe sobre os Planos de Benefícios da Previdência Social e dá outras providências. **Diário Oficial [da] República Federativa do Brasil**, Brasília, DF, 24 jul. 1991. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/L8213cons.htm>. Acesso em: 14 out. 2012.

CHEN, P. P. Entity-Relationship modeling: historical events, future trends, and lessons learned. In: **SOFTWARE pioneers: contributions to software engineering**. 2002. Disponível em: <http://www.csc.lsu.edu/~chen/pdf/Chen_Pioneers.pdf>. Acesso em: 28 out. 2013.

CHEN, P. P. The Entity-Relationship Model - toward a unified view of data. **ACM Transactions on Database Systems**, v. 1, n. 1. Mar. 1976.

DAHLBERG, I. Conceptual definitions for interconcept. **International Classification**, v. 8, n. 1, p. 16-22, 1981.

DAHLBERG, I. Conceptual structures and Systematization. **International Forum on Information and Documentation**, v. 20, n. 3, p. 9-24, 1995.

DAHLBERG, I. Teoria do conceito. **Ci. Inf.**, Rio de Janeiro, v. 7, n. 2, p. 101-107, 1978. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/viewFile/1680/1286>>. Acesso em: 11 maio. 2014.

DAHLBERG, I. Terminological definitions: characteristics and demands. In: **PROBLÈMES de la définition et de la synonymie en terminologie**. Québec: GIRSTERM, 1983. p. 13-51.

DATAPREV. **Dados abertos**. 2012. Disponível em: <<http://api.dataprev.gov.br/revidencia/anuario/2003/acidentes-do-trabalho.xml>>. Acesso em: 15 jan. 2014.

DUBLIN CORE METADATA INITIATIVE. **DCMI metadata terms**. 2012. Disponível em: <<http://dublincore.org/documents/2012/06/14/dcmi-terms/?v=elements#>>. Acesso em: 2 mar. 2014.

HEUSSER, C. A. **Projeto de banco de dados**. Porto Alegre: Sagra, 1998.

HORVATH JÚNIOR, M. **Dicionário analítico da Previdência Social**. São Paulo: Atlas, 2009.

MÉNDEZ, E.; GREENBERG, J. Linked Data for Open vocabularies and HIVE's Global Framework. **El Profesional de la Información**, v. 21, n.3, p. 236-244, 2012. Disponível em: <<http://dx.doi.org/10.3145/epi.2012.may.03>>. Acesso em: 22 maio 2014.

NOLETO, D. A.; BRITO, P. F. RDF e RDF Schema na representação de páginas do portal CEULP/ULBRA. In: ENCONTRO DE ESTUDANTES DE INFORMÁTICA DO TOCANTINS, 5, 2003, Palmas. **Anais...** [Palmas: s.n.], 2003. p. 113-122.

ORGANIZACIÓN INTERNACIONAL DEL TRABAJO. **Tesauro OIT**. 2011. Disponível em: <<http://www.ilo.org/public/spanish/support/lib/tools/aboutthes.htm>>. Acesso em: 2 jun. 2013.

PEREIRA, D. V. **Modelagem e representação semântica de dados governamentais abertos da Previdência Social Brasileira**. 2014. 169 f. Dissertação (Mestrado em Ciência da Informação) – Universidade Federal Fluminense, Niterói, 2014.

PLACES ontology. [2014?]. Disponível em: <<http://purl.org/ontology/places#>>. Acesso em: 3 mar. 2014.

PODER Vocabulary. [2014?]. Disponível em: <<http://dev.poderopedia.com/vocab/schema>>. Acesso em: 2 mar. 2014.

RAMBAYON, A. **Storing RDF data efficiently with E-R modeling**. 2012. Disponível em: <<http://misclassblog.com/database-design-and-development/storing-rdf-data-efficiently-with-e-r-modeling/>>. Acesso em: 2 mar. 2014.

SALES, L. F.; SAYÃO, L. F.; MOTTA, D. F. Modelagem de relações conceituais para a área nuclear. In: SEMINÁRIO DE PESQUISA EM ONTOLOGIAS DO BRASIL, 5, 2012, Recife). **Anais...** Recife: UFPE, 2012.

SCHEMA.ORG. [2014?]. Disponível em: <<http://schema.org/>>. Acesso em: 2 mar. 2014.

SMITH, J. M.; SMITH, D. C. P. Database abstractions: aggregation and generalization. **ACM Transactions on database systems**, v. 2, n. 2, p. 105-133, Jun. 1977.

SPERBERG-MCQUEEN, C. M.; THOMPSON, H. **XML Schema**. 2010. Disponível em: <<http://www.w3.org/XML/Schema>>. Acesso em: 2 mar. 2014.

VANDEBUSSCHE, P.-Y. **Linked Open Vocabularies (LOV)**. [2014?]. Disponível em: <<http://lov.okfn.org/dataset/lov/index.html>>. Acesso em: 2 mar. 2014.

W3C. **The Cambridge comunique**: w3c note 7 October 1999. 1999. Disponível em: <<http://www.w3.org/TR/schema-arch>>. Acesso em: 2 fev. 2014.

WERSIG, G. Information Science: the study of postmodern knowledge usage. **Information Processing & Management**, v. 29, n. 2, p. 229-239, 1993.

XU, L.; LEE, S. W.; KIM, S. E-R Model based RDF data storage in RDB. In: IEEE International Conference on Computer Science and Information Technology (ICCSIT) (3rd., 2010, Chengdu). **Proceedings...** 2010. v. 9. Disponível em: <<http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=5550976>>. Acesso em: 2 mar. 2014.

TEORIA DA TERMINOLOGIA E TEORIA DA ONTOLOGIA: UM COMPARATIVO BASEADO NA CRIAÇÃO DE DEFINIÇÕES PARA TERMOS

THEORY OF TERMINOLOGY AND THEORY OF ONTOLOGY: A COMPARISON BASED ON THE CREATION OF DEFINITIONS FOR TERMS

Elisângela Cristina Aganette
Maurício Barcellos Almeida

Resumo: O presente artigo examina noções da Teoria da Terminologia e da Teoria da Ontologia Aplicada, as quais possibilitam a construção de vocabulários para representar parte da realidade em um domínio do conhecimento. De forma a conduzir esse exame, abordam-se questões advindas das dificuldades em construir vocabulários no domínio da biomedicina, identificadas na literatura e no escopo de um projeto de pesquisa em andamento. Tais questões envolvem, principalmente, a ambiguidade e a dificuldade em formar conceitos e relações entre conceitos de forma consensual. O artigo se propõe explicar como é feita a criação de definições para conceitos de acordo com os preceitos das duas teorias mencionadas, de forma contribuir para a avaliação de quais aspectos de uma das teorias podem ser úteis no contexto da outra, e mesmo porque alguns aspectos de uma teoria não são aplicáveis à outra. Em um cenário que abrange diferentes sistemas de organização do conhecimento, inclusive ontologias, espera-se contribuir para um melhor entendimento da aplicação das duas teorias, bem como fornecer subsídios para auxiliar ao Cientista da Informação no uso das mesmas.

Palavras-chave: Teoria da Terminologia; Teoria da Ontologia Aplicada; Conceitos; Relações entre conceitos.

Abstract: This article assesses some usual notions in the scope of the Theory of Terminology and of the Theory of Applied Ontology. Both of these theories make possible the construction of vocabularies to represent part of the reality in a certain knowledge domain. In order to carry out the assessment, we approach issues arising from the inherent difficulties in building biomedical vocabularies, which can be identified both in the literature and in the scope an ongoing research project. Such issues involve mainly the ambiguity and the difficulty that one faces when trying to consensually create concepts and relations among concepts. This article proposes to explain the process of creation of definition for concepts according to principles of the both the aforementioned theories, in order to contribute to the assessment of which aspects of one of these theories could be useful in the context of the another one, and even why some aspects of a theory are not suitable to be applied in the context of another one. Within a research field that comprehends several different knowledge organization systems, including ontologies, we hope to contribute to a better understanding about the application of both theories, as well as to provide support to Information Scientists in employing them.

Keywords: Theory of Terminology; Theory of Applied Ontology; Relations among concepts.

1 INTRODUÇÃO

A representação da informação e do conhecimento constitui-se atualmente em um campo de investigação desafiante e de importância crucial na sociedade tecnológica. Ao mesmo tempo em que teorias já consagradas encontram aplicações diversas no novo cenário caracterizado pela prevalência de recursos digitais, outras teorias relevantes têm sido

aplicadas apenas recentemente, na busca por formas mais efetivas de representação do conhecimento.

Um exemplo do primeiro tipo de teoria, já há muito adotado na Ciência da Informação, é a amplamente conhecida Teoria da Terminologia. Também denominada Ciência Terminológica, refere-se ao conjunto de princípios que regem o estudo dos termos e suas aplicações na criação de modelos da realidade. O termo “terminologia” apresenta multiplicidade de sentidos, se referindo a: i) uma disciplina, onde a ênfase recai no estudo de termos; ii) uma prática, em que estabelece um conjunto de princípios que regula a coleta e o tratamento de termos; iii) um produto, resultado da prática, que consiste de um conjunto de termos de um assunto especializado mais sua codificação em vocabulários (CABRÉ, 1993).

Um exemplo de teoria aplicada apenas nos últimos anos em Ciência da Informação é a chamada Teoria da Ontologia Aplicada. O termo “ontologia” tem ganhado destaque nos últimos anos em diversas áreas, mesmo que também existam diferentes entendimentos para seu significado. A ontologia como disciplina filosófica, entretanto, a qual estuda o que existe, não é algo novo e teve início ainda com Aristóteles.

O estudo da ontologia tem sido conduzido em duas vertentes principais: i) aquela relacionada ao estudo da lógica e suas aplicações em um ambiente computacional; ii) aquela relacionada ao estudo dos princípios filosóficos capazes de auxiliar a criação de modelos efetivos da realidade. Enquanto o primeiro tipo de estudo tem lugar em campos da Ciência da Computação e gera artefatos computacionais, o segundo tipo – identificado como “Ontologia Aplicada” – encontra aplicação em representação do conhecimento na Ciência da Informação (ALMEIDA, 2013).

A interseção entre as duas teorias (Terminologia e Ontologia) pode ser verificadas especificamente em biologia e medicina. Nesses domínios do conhecimento os trabalhos de Eugen Wüster – considerado o pai da Terminologia – influenciam sobremaneira a criação de vocabulários, uma vez que são à base de normas internacionais ²²amplamente adotadas para esse fim. Por outro lado, os domínios da biologia e medicina são aqueles em que se observa, em todo mundo, o maior número de aplicações diretas de teorias relacionadas à ontologia.

Nesse contexto, caracterizado pelo emprego de abordagens diversas para fins de representação, cabe investigar as similaridades e as diferenças entre as teorias subjacentes. O presente artigo é uma iniciativa nesse sentido conduzida no âmbito de um projeto de

22 ISO/TS 17117:2002 – Health Informatics: Controlled Health Terminology, Structure and High-Level Indicators

construção de uma ontologia biomédica de grande porte (ALMEIDA *et al*, 2011). Apresenta-se aqui uma investigação em curso que confronta noções da Teoria da Terminologia e da Teoria da Ontologia Aplicada. Busca-se verificar em que aspectos essas teorias são próximas (ou não), e se procedimentos da terminologia podem colaborar na construção da ontologia enquanto artefato, a qual é a contribuição do projeto de pesquisa já mencionado.

Para tal, esse artigo apresenta resultados parciais de pesquisa em andamento, a qual compara especificamente a forma como as duas teorias propõem a criação de “definições” para termos no domínio da biomedicina. Espera-se contribuir para um melhor entendimento da aplicação de cada uma das teorias, bem como fornecer subsídios para auxiliar ao Cientista da Informação no uso das mesmas.

O restante do artigo está organizado da seguinte forma: a seção dois descreve algumas noções mais importantes da Teoria da Terminologia, provendo uma visão geral da temática, definições básicas, princípios e fundamentos. Da forma similar, a seção três fornece o mesmo tipo de visão geral de noções da Teoria da Ontologia. A seção quatro apresenta uma comparação preliminar entre as formas de construir definições tanto da Teoria da Ontologia, quanto da Teoria da Terminologia. Finalmente, a seção cinco apresenta considerações finais e perspectivas de trabalhos futuros.

2 TEORIAS DA TERMINOLOGIA

A presente seção apresenta uma visão geral da Teoria da Terminologia (seção 2.1), bem como introduz noções dessa teoria utilizadas para criar definições para um termo (seção 2.2).

2.1 Visão Geral

A Teoria da Terminologia é por vezes definida como um campo de estudos de caráter inter e transdisciplinar, o que a capacita a reunir um conjunto de saberes para fins de apreensão do fenômeno terminológico por excelência – o termo – cuja essência situa-se na representação lexical do conhecimento especializado e na sua divulgação (SANTOS, 2010). Também é definida como uma disciplina teórica e aplicada que tem como elementos centrais os conceitos e termos técnico-científicos. Faz uso da Linguística, da Lógica, da Comunicação, das Ciências Cognitivas, da Ciência da Informação, dentre outras, para descrever e ordenar o conhecimento e sua transferência (KRIEGER, 2004).

Faz-se necessário distinguir as principais teorias terminológicas no contexto da Ciência da Informação, as quais apoiam a construção de linguagens documentárias, enquanto

instrumentos de comunicação de conhecimento, sendo estas: Teoria Geral da Terminologia (TGT) e a Teoria Comunicativa da Terminologia (TCT).

A TGT, também conhecida como Teoria de Terminologia de Wüster, foi formulada a partir da seleção dos termos da engenharia industrial em vista da sua normatização conceitual, linguística e intralinguística com o objetivo de eliminar as ambiguidades que a linguagem natural insere na comunicação profissional, nacional e internacional. Em contextos prescritivos, estruturados e que tem como objetivo a garantia da univocidade comunicativa, a utilidade da TGT é inquestionável. Em situações de comunicação natural, com base social e pretensões identitárias, a TGT não parece ser suficiente (CABRÉ, 1999).

Segundo essa TGT, a terminologia se ocupa dos conceitos de uma língua técnica ou especial, os quais se relacionam entre si como um sistema de conceitos (CAMPOS, 2001). Quando as pessoas criam conceitos a partir de seu entendimento sobre os objetos do mundo, elas estão de fato relacionando, por um lado, indivíduos, e por outro lado, conceitos e objetos ou indivíduos (WÜSTER, 1981).

O termo “indivíduo” representa qualquer objeto físico que existe. As pessoas se lembram dos indivíduos como representações visuais de seu aparato mental e algumas dessas representações são mais abstratas do que apenas uma memória visual de um objeto. Essas representações mais abstratas são chamadas “conceitos individuais”. A outra categoria de conceitos é a categoria dos “conceitos gerais”, os quais são gerados a partir de um processo de abstração (WÜSTER, 1981).

Segundo Cabré (1999) são preceitos da TGT: i) o conceito antecede à expressão; ii) o conhecimento científico é técnico e universalmente uniforme; iii) a estrutura de um domínio especializado é única em todos os grupos e contextos; iv) o conhecimento científico é neutro; v) o uso dos termos normatizados da comunicação profissional deve ser livre de obstáculos; vi) o termo normatizado apresenta características mais significativas para todos os grupos e contextos.

Na TGT, as definições sobre as relações ganham lugar de destaque, o que é evidenciado por Wüester (1979): quando as relações são tratadas em um nível conceitual, elas passam a sere consideradas relações lógicas e ontológicas, havendo uma diferença considerável entre elas. Enquanto as relações lógicas, também chamadas de relações de abstração, resultam forçosamente da própria compreensão dos conceitos, as relações ontológicas nascem do fato de elevarem-se a um nível de abstração as relações que existem na realidade entre os indivíduos.

As oposições e deficiências da teoria wüsteriana, são evidenciadas na TGT, tais como: i) desconsiderar que os termos pertencem à língua natural; ii) não reconhecer que a terminologia é parte da linguagem natural, com todas as consequências dessa implicação; iii) considerar termos como unidades sem interesse sintático, irrelevantes para o discurso; iv) considerar a univocidade e monosemia que contradizem dados extraídos da realidade (CABRÉ, 1999).

Podemos considerar que, diante do exposto, a função da TGT consiste em rotular a denominação dos conceitos da comunicação profissional. Entretanto, essa posição não contempla a dimensão comunicativa dos termos, nem seus aspectos discursivos ou sua projeção gramatical, uma vez que, na TGT supõe-se que os conceitos sejam estáticos, sem variação semântica e com valores pragmáticos, além de não pressupor um estudo da evolução dos conceitos.

Nesse contexto, surgem várias críticas à proposta de Wüster, o conseqüentemente faz-se apresentar uma nova abordagem. Neste caso, por meio da Teoria Comunicativa da Terminologia – TCT, também conhecida como teoria de Cabré. A referida teoria contempla a variação linguística em toda sua dimensão, assume a condição de adequação dos termos, propõe dar conta dos termos como unidades singulares e, às vezes, similares a outras unidades comunicativas, admitindo a variedade conceitual e denominativa, levando em conta a dimensão textual e discursiva dos termos (CABRÉ, 1999).

Um projeto terminológico vinculado teoricamente à TCT deve refletir na sua prática os seguintes pressupostos gerais: i) o objeto central da Terminologia são as unidades terminológicas e não os conceitos. Eleger as unidades como objeto central significa reforçar uma perspectiva linguística e uma abordagem semasiológica; ii) não há uma diferença a priori entre termo e palavra, o que há são signos linguísticos que podem realizar-se no discurso como termo ou palavra dependendo da situação comunicativa; iii) os níveis lexical, morfológico, sintático e textual podem veicular conhecimento especializado; iv) os termos devem ser observados no seu ambiente natural de ocorrência, ou seja, nos discursos especializados; v) a variação conceitual e denominativa deve ser considerada; vi) do ponto de vista cognitivo, as unidades terminológicas: vii) estão subordinadas a um contexto temático; viii) ocupam um lugar preciso num mapa conceitual; ix) o seu significado específico é determinado pelo lugar que ocupam nesse mapa (CABRÉ, 2003).

2.2 Como construir definições para termos de acordo com a Terminologia

Uma questão importante para os objetivos do artigo é verificar quais procedimentos a Teoria da Terminologia sugere para a criação de definições de termos. Apresentam-se nessa seção algumas noções sobre com o assunto é tratado no âmbito da terminologia.

Na teoria terminológica, uma definição é criada a partir de um conceito. As normas terminológicas apontam o conceito como uma unidade subjetiva criada a partir de uma convenção de características. Conceitos são representados pelos termos, os quais são designações verbais. Na *International Organization for Standardization* (ISO), o termo é considerado a unidade mínima da terminologia (ISO-704, 2000).

Uma definição é um enunciado que descreve um conceito permitindo diferenciá-lo de outros conceitos. A definição pode ser formulada de duas maneiras básicas: i) definição por compreensão – definição intensional – a qual compreende a menção ao conceito genérico mais próximo (conceitos superordenados) já definido ou conhecido, mais as características distintivas que delimitam o conceito; ii) definição por extensão – definição extensional – a qual descreve o conceito pela enumeração exhaustiva dos conceitos aos quais se aplica (conceitos subordinados), que correspondem a um critério de divisão (ISO 1087-1, 2000).

Um exemplo de definição intensional é lâmpada incandescente é a lâmpada elétrica cujo filamento é aquecido por uma corrente elétrica de tal modo que ela emite luz; um exemplo de definição extensional é gases nobres são hélio, neônio, argônio, criptônio, xenônio e radônio (ISO 1087-1).

Além disso, uma definição não é exclusiva e pode variar conforme a fonte. Criar uma definição equivale a limitar o termo segundo algum critério, de forma a fixar os limites de um conceito ou ideia. Quando diferentes conceitos têm características semelhantes, é admissível que existam relações entre eles: i) relações hierárquicas, mantidas entre gênero e espécie, espécie e espécie, espécie e indivíduo; ii) relações partitivas, mantidas entre todo e partes, ou entre partes e subpartes; iii) relações da oposição, que expressam contradição; iv) relações funcionais, aplicadas principalmente a conceitos que expressam processos (DAHLBERG, 1978).

3 TEORIA DA ONTOLOGIA APLICADA

A presente seção apresenta uma visão geral da Teoria da Ontologia Aplicada (seção 3.1), bem como noções que tal teoria sugere para criar definições (seção 3.2).

3.1 Visão Geral

Vista como disciplina, a ontologia é a ciência que estuda o que existe, dos tipos das coisas e das estruturas dos objetos, bem como suas propriedades; dos eventos, dos processos e das relações entre as coisas em toda a realidade.

Na literatura da Ciência da Informação, o termo ontologia começou a ser utilizado no final da década de 1990, momento em que instrumentos e métodos de classificação começaram a despertar o interesse também de pesquisadores da Ciência da Computação (VICKERY, 1997). Naquele contexto, a Internet e a recursos digitais exigiam forma de organização da informação até então objeto de estudo da Ciência da Informação.

Uma ontologia, de acordo com Gruber (1993), é uma especificação explícita e formal de uma conceitualização compartilhada. Uma “conceitualização” refere-se a um modelo abstrato de algum fenômeno no mundo que identifica conceitos relevantes daquele fenômeno. O termo “explícito” significa que os tipos de conceitos usados e as restrições a esses conceitos estão definidas explicitamente; assim como “formal” refere-se ao fato de que a ontologia deve ser legível para as máquinas.

No âmbito da área de Ciência da Informação pesquisadores que atuam em organização e representação do conhecimento tem se dedicado cada vez mais ao estudo das ontologias conforme destacam Sales, Campos e Gomes (2006, p. 64): “um importante elemento das ontologias é a representação do conhecimento, que envolve análise semântica, em especial para a organização de sistemas de conhecimento que, nas ontologias, visam aperfeiçoar a recuperação, ou seja, prover o computador de mais inteligência”.

Existem várias descrições sobre o que compõe uma ontologia. A estrutura de uma ontologia pode variar, mas pelo menos alguns componentes são consensuais: conceitos, relações, definições e axiomas. Gómez-Pérez *et al.* (2004) sugerem que uma ontologia é então formada pelos seguintes componentes:

- i) um conjunto e uma hierarquia de classes, as quais são representativas de conceitos; ii) um conjunto de relacionamentos entre classes, por exemplo, o relacionamento entre os conceitos de pessoa e carro é o relacionamento de ser-dono; iii) um conjunto de funções, onde uma função é um caso especial de relacionamento em que um conjunto de elementos tem uma relação única com outro elemento, por exemplo, a função de ser pai-biológico; iv) um conjunto de axiomas, ou seja, declarações sempre verdadeiras, por exemplo, toda pessoa tem uma mãe; v) um conjunto de instâncias, que representam as coisas reais do mundo, agrupadas em classes de acordo com características, por exemplo, João é instância da classe pessoa.

Do ponto de vista dos tipos de ontologias, existem várias classificações propostas. Para Guarino (1998), os dois tipos mais importantes de ontologias são classificadas em: i) ontologias de alto-nível, que descrevem conceitos gerais tipicamente independentes de um

problema particular ou domínio; ii) ontologias de domínio, que descrevem o vocabulário genérico de um domínio, através da especialização de conceitos introduzidos nas ontologias de alto-nível.

Cabe destacar que ontologias de alto-nível são as que possuem maior capacidade de reutilização, uma vez que definem conceitos genéricos; enquanto as ontologias de aplicação são as que possuem menor capacidade de reutilização, por definirem conceitos relativos a uma aplicação específica. As ontologias de alto nível ainda possibilitam uma primeira visão de mundo, que auxilia na tarefa de organizar e representar parte da realidade em estudo.

3.2 Como construir definições para termos de acordo com a Ontologia

A ontologia, ao estudar o que existe e as relações entre o que existe, permite ordenações e classificações. Existem, entretanto, diversas teorias ontológicas, dentre as quais, aquelas consideradas as mais importantes são: a visão aristotélica, que se volta para a estrutura da realidade, e estuda as coisas, e não a forma como nos referimos às coisas; a visão ou kantiana, que se volta para a razão e o pensamento (COCCHIARELLA, 2001). No presente artigo, para estudar a criação de definições, adota-se a visão aristotélica.

O método aristotélico para definir novas categorias baseia-se em noções de *genus* e *differentia*. A *differentia* consiste de propriedades que distinguem diferentes espécies do mesmo gênero. Um exemplo clássico é que todos os homens são mortais, Sócrates é um homem, então Sócrates é mortal (SANTOS, 2010).

As proposições utilizadas por Aristóteles permitem que um raciocínio possa ser transposto para a máquina. As ontologias possuem definições criadas em linguagem lógica, ou seja, passíveis de interpretação por máquinas e não somente por humanos, bem como relações explícitas (BECK; PINTO, 2002).

Uma visão atual baseada em Aristóteles é conhecida como “realismo ontológico” (SMITH; CEUSTERS, 2010). Segundo essa visão, a interpretação de categorias e de relações ontológicas não dependentes da percepção humana. Para se organizar e classificar a realidade do mundo, o ser humano recorre à criação de sistemas de símbolos e, a partir deles, consegue formular representações. No âmbito da ontologia, as definições são relativas a termos, os quais, por sua vez, se referem e, se equivalem, a um objeto do mundo real. Não se consideram conceitos, criações da mente humana, tal como ocorre na terminologia.

Smith (2013) afirma que uma definição é uma declaração de condições necessárias e suficientes. Por exemplo, as seguintes condições são necessárias, bem como suficientes, para x ser um triângulo: x tem três lados, cada lado de x é reto, x é uma figura fechada, x está em

um plano, os lados de x estão ligados por suas pontas. Toda entidade que satisfaz todas essas condições é também um triângulo. Toda entidade que é um triângulo satisfaz a todas essas condições.

De forma geral, uma declaração de condições necessárias e suficientes que envolvem duas entidades A e B, considera: sendo A uma condição necessária para ser um B, então cada B é um A; sendo A uma condição suficiente para ser um B, então cada A é um B; A é uma coisa que satisfaz B.

Outro preceito da ontologia para a criação de definições, apontada por Smith (2013), é o princípio da não circularidade. Uma definição deve ser estabelecida com o uso de termos mais inteligíveis, e não a partir dele mesmo. Por exemplo, definir o termo “hemólise” como “a causa da hemólise”; o “telefone sem fio” como “um telefone que não tem fio”. Além disso, a definição deve ter um formato definido, ou seja, uma entidade S deve ser definida como “S é um G que Ds”, onde G (*genus*) é o termo pai do termo S (espécie) e D é a *differentia*. Por exemplo, ser humano é um animal que é racional (SMITH, 2013).

A principal questão que parece trazer alguma polêmica entre a forma de definir na Teoria da Ontologia e na Teoria da Terminologia é a própria noção de conceito. Não está claro se um conceito é uma entidade mental, uma entidade teórica compartilhada, ou uma entidade linguística (SMITH, 2004). Essa questão é abordada na seção seguinte em mais detalhe, quando são confrontadas as diferentes formas de definir.

4 CONFRONTO DE ASPECTOS DA TEORIA DA ONTOLÓGICA E DA TEORIA DA TERMINOLOGIA

Para confrontar os aspectos das duas teorias em estudo, a pesquisa em andamento avalia a criação de definições no processo de construção de uma ontologia biomédica de grande porte sobre o sangue humano (ALMEIDA *et al.*, 2011). Uma simples busca pelo termo “*blood*” em quatro diferentes sistemas terminológicos (CYC²³, GALEN²⁴, SNOMED²⁵ e UMLS²⁶), obtém definições e classificações bem diferentes para o mesmo conceito.

De fato, exemplos reais da prática médica demonstram as dificuldades e os desafios em criar vocabulários biomédicos consensuais. Exemplos adicionais dessa realidade, além de

²³ <http://www.cyc.com/platform/opencyc>

²⁴ <http://www.opengalen.org>

²⁵ <http://www.ihtsdo.org/snomed-ct/>

²⁶ <http://www.nlm.nih.gov/research/umls/>

considerações teóricas, são apresentados na seção 4.1. Um confronto entre os procedimentos para criação de definições nas duas teorias analisadas é apresentado na seção 4.2.

4.1 Considerações teóricas sobre a comparação entre ontologia e terminologia

No âmbito da biomedicina, onde a investigação em curso nesse artigo tem lugar, existem diferentes formas pelas quais um termo pode ser definido. Liss *et al.* (2003) exemplificam como um termo amplamente utilizado na prática médica corrente é descrito e definido em livros-texto médicos. Do ponto de vista da definição do termo, as definições encontradas em livros pertencem a categorias distintas: i) declarações que servem como critério para definição; ii) declarações que servem como critério para reconhecimento; iii) declarações que servem para evidenciar características; iv) declarações que servem como condições necessária ou suficiente.

É possível distinguir pelo menos duas implicações dessas diferentes interpretações identificadas. A primeira é uma questão de linguagem, especificamente de má comunicação, que ocorre quando dois especialistas com conhecimento médico suficiente usam diferentes termos para se referir as mesmas coisas. A segunda é uma questão de ontologia, ou seja, qual das definições representa a forma mais correta para descrever as coisas do mundo.

No segundo caso, a questão da ontologia, as diferenças não ocorrem simplesmente pelo mero uso de termos diferentes. Os diversos especialistas envolvidos acreditam em “diferentes ontologias” a respeito das entidades em questão. Ter diferentes ontologias não diz respeito a usar diferentes termos para descrever o mundo, mas sim a ter diferentes visões de como o mundo é constituído.

Para criar um modelo de diagnóstico que seja útil na prática médica é importante buscar uma visão ontológica consensual. Essa situação não pode ser resolvida via linguagem. Conforme explica Grenon (2008), a linguagem natural pode servir como uma pista para o ontologista, mas não deve ser o critério de corretude dos resultados finais de seu trabalho.

Do ponto de vista da ontologia, que permitiria criar um modelo de diagnóstico adequado, o problema parece residir na imprecisão e ambiguidade na definição do termo “conceito”, o qual é de fundamental importância no contexto da terminologia.

O termo “conceito” pode ser interpretado de acordo com duas diferentes visões: i) pelo viés realista, no qual conceitos são entidades independentes dos seres humanos e que podem ser descobertos por eles; ii) pelo viés construtivista, que advoga que conceitos são criados pelos seres humanos e, portanto dependentes da mente. Wüster (1979), por exemplo, parece

aderir à caracterização construtivista, enquanto as principais teorias contemporâneas da ontologia como a de Smith (2014), parecem aderir ao viés realista.

Berzell (2010) explica que as ambiguidades presentes, por exemplo, nas teorias de Wüster, impossibilitam seu uso para a criação de modelos de diagnóstico compatíveis com a realidade médica. Em algumas passagens da obra de Wüster, conceitos são referenciados como entidades individuais, mentais e, portanto subjetivas; em outras passagens, conceitos são referenciados como entidades objetivas (SMITH *et al.*, 2005).

Esse tipo de ambiguidade impacta, em última instância, na normalização da ISO, uma vez que Wüster foi um dos principais fundador e mantenedores do comitê técnico de terminologia. A ISO 704 (2000), por exemplo, estabelece que conceitos não devem ser confundidos com objetos abstratos ou imaginados (isto é, objetos concretos, abstratos, ou imaginados em um dado contexto são observados e conceitualizados mentalmente e então, uma designação é atribuída ao conceito e não ao objeto em si. Para esse padrão internacional, a conexão entre objeto e sua designação ou definição é feita por meio do conceito, um nível de abstração mais alto.

Uma conclusão relevante para os objetivos da pesquisa em andamento pode ser obtida desse breve comparativo. Os fundamentos da Teoria da Terminologia são claramente baseados em teorias da linguagem, onde se busca o entendimento de como usar conceitos ou pensamentos como uma forma de comunicar o que os signos fazem referência. Então, as ontologias não são equivalentes às terminologias como sugere a ISO/TS 17117:2002 (*Health Informatics – Controlled Health Terminology, Structure and High-Level Indicators*) criada para a área médica.

A visão de conceitos como pensamentos não está alinhada com o fato de que ontologias são artefatos representacionais que se referem a universais da realidade e as relações entre eles. Se o que se deseja é criar teorias bem fundamentadas da realidade, o que, conforme já mencionado, se se faz necessário no âmbito da medicina e biologia, as ontologias são o artefato a considerar (MENDONÇA; ALMEIDA, 2012). Caso se adotem as abordagens da terminologia, não existe garantia que o resultado obtido esteja alinhado com as melhores teorias científicas disponíveis.

4.2 Considerações práticas sobre a comparação entre ontologia e terminologia

Uma vez apresentadas considerações teóricas que revelam importantes diferenças e similaridades entre as teorias em estudo, segue uma caracterização de cunho prático, uma lista dos procedimentos necessários para criar definições de acordo com as duas teorias. Esses

procedimentos têm sido usados no âmbito de um projeto de biomedicina (ALMEIDA *et al.*, 2011) para comparar definições existentes e criar novas definições para uma ontologia no domínio da hematologia. Espera-se com essa caracterização prática evidenciar mais diferenças entre as teorias da terminologia e da ontologia.

I) Procedimentos para definir termos a partir da Teoria da Terminologia (de a até g)

a) Separar o termo. Exemplo: “*mechanical mouse*”.

b) Obter uma elucidação preliminar sobre o significado do termo em alguma fonte (dicionário etc.). Exemplo: “<computing> computer mouse in which movements are detected by a ball on its underside that activates rollers in physical contact with the ball”).

c) Estabelecer o conceito superordenado imediatamente superior no contexto de uso do termo. Exemplo: “*computer mouse*”.

d) Estabelecer as características que delimitam o termo, a partir de características que diferenciam o conceito dados dos conceitos coordenados;

e) Formular e escrever a primeira versão da definição de forma sistêmica:

e1) A definição deve estar na forma de uma declaração (não necessariamente uma sentença);

e2) Se o termo a ser definido é um substantivo, a definição consiste da combinação de:

i) o substantivo; (exemplo: “*mechanical mouse*”)

ii) o verbo “ser” indicando a copula sujeito-predicado; (é um)

iii) conceito super-ordenado (exemplo: “*computer mouse*”)

iv) delimitadores que indicam as características que delimitam os conceitos sob definição (no qual, que, etc)

Exemplo: “[A] mechanical mouse [is a] computer mouse in which movements are detected by a ball on its underside that activates rollers in physical contact with the ball”).

f) Verificar deficiências na primeira versão da definição

f1) Verificar princípio da não circularidade (dentro do mesmo conceito)

i) Exemplo: errada, circular = “tree height is a tree height measured from the ground surface to the top of a tree”;

ii) Exemplo: correta, não circular: “tree height is the distance between the ground surface and the top of a tree”

f2) verificar princípio da substituição (dentro de sistemas de conceitos)

i) Exemplo (circular): “haploid life cycle is the definition period in an organism's life involving one generation when only the multicellular stage is haploid”;

ii) Exemplo: correto se é possível substituir “*haploid*” (circular) por uma definição em separado, por exemplo, “*haploid = having a single set of chromosomes in the nucleus of each cell*”.

f3) Verificar se a definição é precisa

f4) Verificar e eliminar definições negativas

i) Exemplo: correta, afirmativa = “deciduous tree is a tree that loses its foliage seasonally”;

ii) Errada, negativa = “deciduous tree is a tree other than an evergreen tree”

g) Refinar e propor uma versão final da definição

II) Procedimentos para definir termos a partir da Teoria da Ontologia (de a até g)

a) Separar o termo - Exemplo: “*plasma membrane*”

b) Obter uma elucidação preliminar sobre o significado do termo em alguma fonte (dicionário, etc.) - Exemplo: “*the semipermeable membrane enclosing the cytoplasm of a cell*”

c) Estabelecer o genus superior no contexto de uso do termo - Exemplo: “*plasma membrane is a membrane*”

d) Estabelecer a(s) característica(s) essencial(is) das coisas (sentido aristotélico), distinguindo o genus de as espécies

i) Para Aristóteles, apenas um desses exemplos pode estar correto:

-“Man is featherless biped”

-“Man is a rational animal”

ii) Para Aristóteles, a essência das coisas deve ser encontrada

iii) Para se encontrar a essência das coisas é preciso fazer ciência

e) Formular e escrever a primeira versão da definição:

i) As definições devem estar na forma: S = Def. um G o qual Ds

-onde “G” (para: genus) é o termo pai de “S”

-onde “S” (para: espécies) na ontologia de referência correspondente

-onde S e G são tipos

Exemplo:

-“human being = def. an animal which is rational”

-“human being = def. intersection of (animal) and (rational)”

f) Verificar se a definição é uma declaração de condições necessárias e suficientes

i) Uma definição deve ser uma declaração de condições necessárias e suficientes

ii) Ser um A é condição necessária para ser um B, então cada B é um A

iii) Ser um A é condição suficiente para ser um B, então cada A é um B

Exemplos:

i) As condições necessárias e suficientes para que X seja um triângulo são:

-X tem exatamente três lados

-cada um dos lados de X é reto

-X é uma figura fechada

-X está em um plano

-os lados de X se juntam em seus finais

-Cada coisa que satisfaz todas essas condições é também um triângulo;

-Cada coisa a qual é um triângulo satisfaz a todas essas condições.

ii) “Completing all of your requirements is both a necessary and sufficient condition for earning your degree. Without completing all requirements, it is impossible to earn a degree, and completing all requirements guarantees earning a degree.”

g) Verificar deficiências na primeira versão da definição

g1) Verificar princípio da não circularidade

i) A definição de um dado termo deve usar termos que são mais inteligíveis, e de mais fácil compreensão. Exemplo (circular): “*hemolysis =def. the causes of hemolysis*”

g2) Verificar e eliminar herança múltipla

Exemplo: “blue thing is a blue thing and car is a thing” + “blue car is a car and blue car is a car”

g3) Verificar princípio da substituição

i) Substituição e a chave para entender definições: se uma definição está correta, então é possível substituir a definição pelo termo definido em todos os contextos e preservar a verdade;

Exemplo:

Nas sentenças:

-cell = def. an anatomical structure which consists of cytoplasm surrounded by a plasma membrane;

-plasma membrane =def. a cell part that surrounds the cytoplasm

A definição de “*plasma membrane*” da segunda sentença pode ser usada para substituir o termo na primeira sentença sem perda de significado.

g4) Verificar princípio do desdobramento

Definições são a principio não elimináveis: pode-se desdobrar os textos que as contém completamente, sem alteração do valor verdade.

h) Refinar e propor uma versão final.

Finalmente, o QUADRO 1 resume algumas características do processo de criar definições nas duas teorias sob estudo.

QUADRO 1 – Exemplo do tipo de comparação pesquisada

Característica básica	Formas de definir entidades	
	Método Aristotélico	Método Terminológico
Condições necessárias e suficientes	Sim	Não
Noção intencional e extensional	Sim	Sim
Não circularidade	Sim	Sim
Exaustividade	Sim	Não

Fonte: Do autor.

5 CONSIDERAÇÕES GERAIS

O presente artigo buscou descrever importantes noções da Teoria da Terminologia e da Teoria da Ontologia Aplicada, por meio de suas definições, princípios e fundamentos. Em seguida, apresentou uma caracterização de definições de acordo com distintas visões das referidas teorias e seus respectivos detalhamentos.

A discussão conduzida no artigo faz parte de uma investigação em curso que objetiva confrontar noções da Teoria da Terminologia e da Teoria da Ontologia Aplicada, verificar em que aspectos essas teorias são próximas (ou não), e se procedimentos da terminologia podem colaborar na construção da ontologia como advoga a norma baseada em princípios da terminologia criada para a área médica (ISO/TS 17117:2002). Examinaram-se noções das duas teorias, em virtude de questões advindas das dificuldades em construir vocabulários no domínio da biomedicina, principalmente no que diz respeito à ambiguidade e a dificuldade em formar conceitos e relações entre conceitos de forma consensual.

Uma possível diferença entre as duas teorias identificadas neste estudo está justamente no objetivo de cada uma delas: enquanto a Teoria da Terminologia visa à recuperação de informação a partir dos conceitos presentes em documentos, a Teoria da Ontologia Aplicada busca criar modelos em um dado domínio da realidade.

Outra diferença observada entre as referidas teorias está na maneira de se criar definições para um termo. Ao se comparar especificamente a forma como as teorias propõem a criação de definições, verificaram-se diferenças. Por um lado, a Teoria da Terminologia afirma que os termos são definidos por meio de conceitos, ou seja, criações da mente humana e assim carregadas de subjetividade; por outro lado, a Teoria da Ontologia Aplicada reafirma e mantém o compromisso mínimo com as entidades do mundo real, buscando ao máximo uma representação fidedigna da realidade. Para ter acesso às entidades do mundo real, a ontologia propõe fazer ciência (SMITH, 2008).

Diante do exposto, pode-se considerar que a teoria terminológica com sua visão orientada ao conceito pode não ser adequada para domínios específicos, como no caso domínios biomédicos. Isso se justifica, não por deficiências teóricas, mas apenas pela utilização e aplicação em domínios sensíveis que exigem a criação de modelos para a prática clínica. Para estes domínios, faz-se necessário lidar com a realidade, com intuito de se descrever e diferenciar os objetos e seus relacionamentos do mundo, ou seja, não apenas conceitos, e tratar as relações entre os objetos independentes da mente humana de acordo com a realidade. Para Smith (2003) trata-se de um “compromisso ontológico” evidenciar as diferenças entre as coisas e os processos que se está tentando representar.

A principal polêmica entre a forma de definir de acordo com as respectivas teorias é agravada pela inexistência de um consenso do que seria um “conceito”. O trabalho terminológico inicia-se com a noção de conceito, a qual possui uma unidade de denominação que é o termo. Assim, um termo designa um conceito que, por sua vez, é o significado do termo. Wüster defende a visão psicológica (conceitos são entidades mentais) e, nesse sentido, nosso conhecimento sobre conceitos está enraizado em experiências da infância, que se encontra disperso por uma coleção de impressões sensoriais diversas (SMITH, 2005). Não está claro, assim, se um conceito é uma entidade mental, uma entidade teórica compartilhada, ou uma entidade linguística (SMITH, 2004).

Para Wüster (1979), antes de atribuir um termo a um conceito, é preciso primeiro “delinear” o conceito, ideia esta que enfrenta sérias objeções, uma vez que para delinear, é preciso listar a totalidade das “características” que formam o conteúdo ou intensão de um conceito. Wüster fornece esclarecimentos conflitantes sobre o que essas “características” podem ser, concebendo-as, por vezes, como se fossem elas mesmas conceitos (de modo que, como outros conceitos, elas existem na mente das pessoas), e em outros momentos como propriedades de objetos existentes no mundo.

Embora as terminologias possam ser empregadas com êxito na representação de significados abstratos, nem sempre elas têm sido consideradas suficientemente precisas para aplicações e usos computacionais (FREITAS; SCHULZ, 2009). Nesse artigo, avaliou-se o aspecto da criação de definições nas duas teorias, e conclui-se que a forma de definir da Teoria da Terminologia pode auxiliar o início do trabalho de construir definições para ontologias, mas não atende a todos os critérios da Teoria da Ontologia Aplicada.

No entanto, não se pode deixar de salientar e reafirmar que as ideias difundidas pela teoria geral da terminologia de Wüester foram de inquestionável importância e influência para os trabalhos terminológicos seguintes. Por meio da definição de conceitos e de princípios para a criação de novos termos, foi possível a comunicação entre os especialistas de diversos domínios do conhecimento. A TGT foi o cerne de todos os trabalhos atualmente realizados para a representação da informação na Ciência da Informação.

Em um cenário que abrange atualmente diferentes sistemas de organização do conhecimento, inclusive ontologias, espera-se que este artigo tenha contribuído para um melhor entendimento da aplicação de cada uma das teorias, bem como fornecer subsídios para auxiliar ao Cientista da Informação no uso das mesmas.

Como trabalhos futuros e na continuidade da pesquisa, a investigação será comparar a forma de construir relações entre entidades de acordo com ontologia e com terminologia. No âmbito da Teoria da Ontologia, já se sabe, só existem relações lógicas, enquanto que para a Teoria da Terminologia as relações são criadas entre conceitos.

REFERÊNCIAS

- ALMEIDA, M. B.; BAX, M. P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Ciência da Informação**, v. 32, n. 3, p. 7-20, 2003.
- ALMEIDA, M.B. Revisiting Ontologies: A Necessary Clarification. **Journal of the American Society of Information Science and Technology**. v. 64, n. 8. p. 1682-1693, 2013. DOI: 10.1002/asi.22861.
- ALMEIDA, M.B.; PROIETTI, A.B.; AI, J.; SMITH, B. The Blood Ontology: an ontology in the domain of hematology. **Proceedings of the International Conference of Biomedical Ontologies**, 2011, US.
- BECK, H.; PINTO, H. S. Overview of approach, methodologies, standards, and tools for ontologies. **The Agricultural Ontology Service (UN FAO)**, Sections 1, 4, and 6, 2002.
- BERZELL, M. **Electronic Healthcare Ontologies: philosophy, the real world and IT structures**. PhD thesis, Linköping University, Faculty of Arts and Sciences, Division of Health and Society, Department of Medical and Health Science. Linköping, Sweden, 2010. 163 p.

CABRÉ, M. T. La terminología. Teoría, métodos, aplicaciones. Barcelona: Antártida, 1993.

CABRÉ, M. T. **La terminología: representación y comunicación**. Elementos para una teoría de base comunicativa y otros artículos. Iula/Universitat Pompeu Fabra, Barcelona, 1999, 369 p.

CABRÉ, M. T. Theories of terminology: their description, prescription and explanation. **Terminology**, v. 9, n. 2, p. 163-200, 2003.

CAMPOS, M. L. A. **Linguagem Documentária: teorias que fundamentam sua elaboração**. Niterói-RJ: EUFF, 2001.

COCCHIARELLA, N. B. Logic and ontology. **Axiomathes**, 12, p. 117-150, 2001.

DAHLBERG, I. Teoria do conceito. **Ciência da Informação**, Rio de Janeiro, v. 7, n. 2, p. 101-107, jul./dez. 1978.

FREITAS, F.; SCHULZ, S. Pesquisa de terminologias e ontologias atuais em biologia e medicina. **RECIIS – R. Eletr. de Com. Inf. Inov. Saúde**. Rio de Janeiro, v.3, n.1, p.8-20, mar, 2009. Disponível em: <<http://ontology.buffalo.edu/bfo/BeyondConcepts.pdf>>. Acesso em: 5 mai. 2014.

GÓMEZ-PÉREZ, A.; FERNÁNDEZ-LÓPEZ, M.; CORCHO, O. Ontological Engineering with examples from the áreas of knowledge Management, e-Commerce and the Semantic Web. Springer-Verlag. London Limited. 2nd printing, 2004.

GRENON, P. A Primer on Knowledge Representation and Ontological Engineering. In MUNN, K. & B. SMITH (Ed.). 2008.

GRUBER, T. R. **Towards principles for the design of ontologies used for knowledge sharing**. Stanford Knowledge Systems Laboratory, Palo Alto, CA, 1993.

GUARINO, N. Formal Ontology in Information Systems. **Proceedings of FOIS'98**, Trento, Italy, 6-8 June 1998. Amsterdam, IOS Press, p. 3-15, 1998.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **Principles and methods of terminology: 704**. 2. ed. [S. l.]: [s. n.], 2000.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. **Terminology work – principles and methods: 1087-1**. Part 1: theory and application = Travaux terminologiques – vocabulaire. Partie 1: théorie et application. [S. l.]: [s. n.], 2000.

KRIEGER, M. G.; FINATTO, M. J. B. **Introdução à Terminologia**. São Paulo: Contexto, 2004.

LISS P-E, ASPEVALL O, KARLSSON D, FORSUM U.: 2003, 'Terms used to describe urinary tract infections – the importance of conceptual clarification.' **APMIS**, v. 111, p. 291-299, 2003

MENDONÇA, F. M.; ALMEIDA, M. B. Modelos e teorias para a representação: uma teoria ontológica sobre o sangue humano. **Anais do XIII Enancib**. Rio de Janeiro: ENANCIB, 2012.

SALES, L. F.; CAMPOS, M. L. A.; GOMES, H. E. Ontologias de domínio: um estudo das relações conceituais e sua aplicação. In: ENCONTRO NACIONAL DE PESQUISA EM

CIÊNCIA DA INFORMAÇÃO - ENANCIB, 7. 2006, Salvador. **Anais...** Salvador: ANCIB, 2006. Disponível em:
<<http://www.marilia.unesp.br/sistemas/enancib/viewpaper.php?id=205>>. Acesso em: maio 2014.

SANTOS, C. **Terminologia e ontologias: metodologias para representação do conhecimento**. Dissertação de Doutorado, Universidade de Aveiro, Aveiro, 2010.

SMITH, B. **Introduction to the Logic of Definitions**. 2013. Disponível em:
<<http://www.slideshare.net/BarrySmith3/logic-ofdefinitions>>. Acesso em: 12 maio 2014.

SMITH, B. **Beyond concepts: ontology as reality representation**. 2004. Disponível em:
<<http://www.ontology.buffalo.edu/ontology>> Acesso em: 10 jun. 2014.

SMITH, B. **Ontology and Informations Systems**. 2003. Available from Internet:
<<http://www.ontology.buffalo.edu/ontology>>. Access: 22 jun. 2014.

SMITH, B. **Video: how to build an ontology**. Disponível em:
<<http://ontology.buffalo.edu/smith/articles/ontologies.htm>> Acesso em: 16 jun. 2014.

SMITH, B.; CEUSTERS, W.; TEMMERMAN, R. Wüsteria. **Proceedings of Medical Informatics Europe**. 2005. Disponível em:
<<http://ontology.buffalo.edu/medo/Wuesteria.pdf>>.

SMITH, B.; CEUSTERS, W. Ontological realism: a methodology for coordinated evolution of scientific ontologies. **Applied Ontology**, IO Press, v. 5, p. 139–188, 2010.

SMITH, B.; MUNN, K. **Applied Ontology: an introduction**. 2008. Disponível em:
<<http://ontology.buffalo.edu/AppliedOntology.pdf>> Acesso em: 20 abr. 2011.

SMITH, B.; WILLIAMS, J.; SCHULZE-KREMER, S. The ontology of the Gene Ontology. In: **Proc. of the AMIA Symposium 2003**. [S.l.]: [s.n.], , 2003. p. 609-613.

SMITH, B. (2003). **Ontology: an introduction**. Longer draft, ontology and information systems. Disponível em: <<http://ontology.buffalo.edu/smith/articles/ontologies.htm>>. Acesso em: 16 jun. 2014.

VICKERY, B. C. Ontologies. **Journal of Information Science**, v. 23, n. 4, p. 277-286, 1997.

WÜESTER, E. Introduction to the general theory of terminology and terminological lexicography. Viena: Springer, 1979.

WÜSTER, E. L' étude scientifique générale de la terminologie, zone frontalière entre la linguistique, la logique, l' ontologie, L' informatique et les sciences des chose. In: RONDEAU, G.; FELBER, E. (Org.). **Textes choisis de terminologie**. Québec: GIRSTERM, 1981. (Fondements théoriques de la terminologie, v. I). p. 57-114.

**A SEMIÓTICA COMO SUBSÍDIO PARA A REPRESENTAÇÃO DO
CONHECIMENTO: UMA ANÁLISE CONCEITUAL SOBRE O TEMA**

**THE SEMIOTICS OF CULTURE IN SOCIOCULTURAL APPROACHES TO
KNOWLEDGE ORGANIZATION: A THEORETICAL AND CONCEPTUAL
ANALYSIS**

Isadora Victorino Evangelista
José Augusto Chaves Guimarães
Carlos Cândido Almeida

Resumo: Em que pese a natureza mediadora da representação do conhecimento, ao estabelecer pontes entre um conhecimento produzido e materializado pela sociedade e a apropriação desse conhecimento para geração de um novo, foi-se visto, por muito tempo, apenas como reflexo de uma prática, em que os manuais de ensino e pesquisa, via de regra, apenas sugerem as partes de maior conteúdo informacional onde podem ser encontrados os assuntos, mas não consideram outros fatores, como a experiência do indexador, suas visões de mundo e qual a melhor maneira de conduzir essa representação. A vista disso, essa área de estudos se volta, hoje, para a necessidade de um aprofundamento teórico e metodológico. Desse modo, e considerando a Semiótica como campo que estuda a representação da realidade por meio de signos, a partir das experiências vividas pelo indivíduo, analisa-se em que medida pode contribuir para a fundamentação teórica e metodológica da representação do conhecimento no intuito de demonstrar como suas representações de realidade podem auxiliar o profissional bibliotecário nas atividades de representação do conhecimento, mais especificamente no que se refere à representação temática ou de conteúdo. Para tanto, e partindo do resgate teórico sobre Representação do Conhecimento e sobre Semiótica, analisa-se a contribuição dos teóricos da área que vem tradicionalmente se dedicando a esse tema (Lara e Fujita, No Brasil, e Mai e Thellefsen, na Dinamarca) relativamente à contribuição teórico-metodológica da Semiótica para a representação de assunto, concluindo-se pela iminência de maiores reflexões teóricas sobre essa questão, de modo a que os processos de indexação possam ser encarados antes como processos interpretativos do que como fazeres técnicos do bibliotecário.

Palavras-chave: Organização do conhecimento. Indexação. Semiótica

Abstract: On the mediating nature of knowledge representation, to build bridges between knowledge produced and materialized by society and the appropriation of this knowledge to a new generation, was seen, for a long time, just as a reflection of a practice in that textbooks and research, as a rule, only suggest higher information content which can be found subjects, but do not consider other factors such as the experience of the index, their worldviews and how best to conduct this representation. As a result, this area of study turns today to the need for a theoretical and methodological deepening. Therefore, and considering the semiotics as a field that studies the representation of reality by means of signs, from the experiences of the individual, is analyzed to what extent can contribute to theoretical and methodological foundation of knowledge representation in order to demonstrate how their representations of reality may help the professional librarian in the representation of knowledge activities, specifically with regard to the subject or content representation. To this end, and based on the theoretical survey on Knowledge Representation and Semiotics, analyzes the theoretical contribution of the area that has traditionally been dedicated to this topic (Lara and Fujita, in Brazil, and Mai and Thellefsen, Denmark) for the theoretical and methodological contribution of semiotics to the representation of the subject, concluding the brink of major theoretical

reflections on this issue, so that the indexing processes can be seen as interpretative rather than as technical processes doings librarian.

Keywords: Knowledge organization. Indexing. Semiotics.

1 INTRODUÇÃO

A organização do conhecimento, enquanto campo de estudos de natureza teórica e metodológica apresenta uma dimensão natureza mediadora entre um conhecimento produzido e materializado pela sociedade e a apropriação desse conhecimento para geração de um novo. Nesse contexto, um conjunto de processos é levado a cabo, dentre os quais aquele que, valendo-se de um conjunto de instrumentos, denominados linguagens de indexação, chega à geração de produtos como catálogos e índices, viabilizando o acesso, o uso e a apropriação da informação. A esse processo, a Ciência da Informação denomina como representação do conhecimento.

A área de representação do conhecimento e da informação – entendidas como a extração de assuntos através de termos para posterior recuperação –, vista por muito tempo apenas como reflexo de uma prática, em que os manuais de ensino e pesquisa – como por exemplo as normas técnicas ou a classificação de Dewey, via de regra, apenas sugerem as partes de maior conteúdo informacional onde podem ser encontrados os assuntos, mas não consideram outros fatores, como a experiência do indexador, suas visões de mundo e qual a melhor maneira de conduzir essa representação. A vista disso, essa área de estudos sobre aspectos linguístico-semióticos se volta, hoje, para a necessidade de um aprofundamento teórico e metodológico, aspecto que vem sendo objeto de estudo de autores Lara (1993a, 1993b, 2001, 2002, 2006), Lara e Tálamo (2007), Tálamo (2001, 2009) Tálamo e Lara (2006, 2009) e outros. Desse modo, e considerando a Semiótica como campo que estuda a representação por meio de signos (SANTAELLA, 2005), questiona-se em que medida esta pode contribuir para a fundamentação teórica e metodológica da representação do conhecimento e da informação. Objetiva-se, portanto, levantar apontamentos semióticos para compreender a prática e os processos de organização do conhecimento a partir dos afazeres do profissional.

Assim, busca-se apresentar, por meio de uma revisão conceitual, a contribuição dos teóricos da área, e em que medida sua abordagem se torna inerente ao processo de representação do conhecimento e da informação, em especial no âmbito temático. Para tanto, apresenta-se, em um primeiro momento, uma seção sobre representação do conhecimento, a que se segue uma breve exposição da escola de Semiótica. Em seguida, aborda-se como a Semiótica pode se relacionar com a representação de assunto, a partir das discussões ocorridas

no cenário nacional e internacional sobre essas questões, para que se possa chegar a conclusões a respeito.

2 A REPRESENTAÇÃO DO CONHECIMENTO E REPRESENTAÇÃO DA INFORMAÇÃO

Como destaca Barité (2001), a Ciência da Informação centra sua ênfase na gestão, no tratamento e no uso social da informação. Mais especificamente no que se refere ao tratamento informacional, tem-se a Organização do Conhecimento, como matriz teórico-metodológica que subsidia as práticas de Organização da Informação, voltadas a realidades documentais específicas.

Para Hjørland (2008), a Organização do Conhecimento tem por objeto as questões e procedimentos ligados ao tratamento informacional, atentando-se para a natureza e a qualidade do conhecimento, de tal forma que se possa propiciar a utilização eficaz de sistemas para organizar documentos e suas representações. Para o autor, a Organização do Conhecimento classifica-se como sendo a base teórica para a Organização da Informação. Smiraglia (2013) ressalta ainda o fato de a Organização do Conhecimento dedicar-se à ordem conceitual do conhecimento, onde são estudadas as heurísticas de ordenação do conhecimento, de modo a que se possa ter soluções para eventuais problemas trazidos pela Organização da Informação.

Trabalha-se, aqui, com um conhecimento que, uma vez registrado e socializado (ou, em outras palavras, materializado em um documento), insere-se em um contexto onde se fazem necessários instrumentos para torna-lo disponível, para que possa subsidiar a produção de um novo conhecimento, em um movimento helicoidal (GUIMARÃES, 2008). Nesse contexto, duas abordagens se apresentam e se completam: a abordagem descritiva, que tem por objeto a representação dos aspectos físicos do documento, com vistas a sua localização, e a abordagem temática, que resulta na representação do conteúdo informacional do documento, valendo-se de procedimentos ligados à análise documental. Como se pode observar, a representação do conhecimento atua como resultado de um processo de análise inerente à organização do conhecimento, quando se busca estabelecer equivalências por meio de termos que possam expressar o conteúdo documental (ALMEIDA e GUIMARÃES, 2008).

Nessa etapa final – a representação - o documento passa por um processo de interpretação por parte do bibliotecário, de tal forma que se possa estabelecer uma ponte entre o recurso e o usuário, ponte essa que se materializa por meio de termos estabelecidos pelo bibliotecário, que passam a atuar como representações de realidade daquele documento

(GUIMARÃES, 2008).

De acordo com Dahlberg (2006), representação do conhecimento vai além de uma lógica de representação conceitual, atingindo “todas as questões relacionadas a nomear conceitos pelos termos mais aptos, onde as questões de terminologia também devem ser consideradas.” (DAHLBERG, 2000, p. 12).

Como instrumentos para a Representação do Conhecimento temos aquilo que a corrente anglo-saxônica denomina linguagens de indexação e a corrente francesa denomina linguagens documentais, que se caracterizam por serem elementos criados artificialmente “que auxiliam os processos de tratamento e recuperação de documentos, sendo utilizadas em etapa subsequente à identificação e à seleção de conceitos contidos nos conteúdos documentais” (SALES et al., 2011, p. 3).

As referidas linguagens são constituídas a partir de termos, notações ou símbolos, construídos artificialmente, e visam ao tratamento e à recuperação da informação, podendo diferir tipologicamente a partir da maneira como são estruturadas como por ex. sistemas de classificação, listas de cabeçalhos, tesauros, taxonomias, ontologias, etc. (SALES et. al., 2011, p. 9).

A representação do conhecimento e da informação está, portanto, comprometida com a Organização do Conhecimento, se estabelecendo através de âmbitos como a classificação e a indexação.

3 SEMIÓTICA PEIRCEANA

Para que se possa abordar, ainda que sucintamente, a Semiótica como universo de estudos, necessário se faz recorrer, preliminarmente, à Semiologia, explicitando a diferença entre os termos, como área que também busca centrar suas reflexões nos signos linguísticos, mas mais especificamente a partir das estruturas da linguagem. É importante explicitar no entanto, que é impossível resumir em poucas páginas os conceitos das duas escolas.

A Semiologia, de origem europeia encontra bases na Linguística, a partir dos estudos da linguagem e da comunicação por meio de signos, e nasce a partir da necessidade de que, em algum momento, a realidade seja representada. Para a Semiologia, nada que não possa ser expresso por meio de palavras, é considerado. Portanto, as experiências e a cognição humanas não são cogitadas, pois linguisticamente elas não podem ser representadas (NOTH, 1996).

De acordo com Noth (1996), Saussure apresenta a linguagem como a conjunção de LINGUA + FALA em que a linguagem atua como um conjunto de regras, normas e palavras, a língua, como um sistema subjacente, e fala, como um mecanismo de uso. Como destaca o

referido autor, esse campo encontrou significativa contribuição de teóricos como Hjelmslev, que se preocupou em formalizar a teoria linguística através de sua ideia sobre signo glossemático; Barthes, que apresentou os conceitos de denotação e conotação como elementos da Semiologia; Morris, que dividiu a Semiologia em sintática, semântica e pragmática consideradas os fundamentos para a teoria dos signos; Greimas, com seu conceito de Semiologia discursiva, e Eco, ao propor um campo/espço de estudos semióticos que fosse além de uma simples disciplina. Nesse contexto, têm lugar, também, as concepções de Semiótica de Peirce, que se desenvolveram notadamente nos Estados Unidos e constituíram uma escola à parte.

A denominada Semiótica Peirceana, a partir das concepções de Charles Sanders Peirce, encontrou fundamento nas teorias científicas, em que todas as conclusões necessitam estar fundamentadas a partir de evidências, que futuramente serão interpretadas e estarão sujeitas às críticas da comunidade científica. Sendo assim, nenhuma teoria científica é infalível, devendo estar sempre submetidas a correções e revisões (SANTAELLA, 2005).

Considerando também as experiências e outras formas de comunicação que não sejam estritamente humanas, como a linguagem animal, a Semiótica vai além das preocupações da Linguística, sobre a lacuna entre a percepção e a cognição, para abordar também os extremos desses dois, considerados extralinguísticos. Assim, qualquer coisa que substitui/representa outra pode ser considerada um signo, como por exemplo, uma palavra que denomina um objeto; ela não é o objeto em si, mas apenas uma representação da realidade do que aquele objeto é (SANTAELLA, 2005).

O signo, para Peirce, é composto por uma relação triádica indissolúvel, composta pelo seu fundamento, seu objeto e seu interpretante, sendo que não há pensamento sem o signo pois a mente humana está o tempo todo atribuindo signos para coisas reais, mesmo que não haja plena consciência desse fator. Para tanto, três categorias se apresentam: a primeiridade, quando ainda não se tem a consciência do que se está ao redor (como, p.ex., a não-consciência de todos os elementos quando se está na estrada, como a cana-de-açúcar se movendo na plantação); a segundidade, quando se toma consciência de que não se está sozinho em um determinado espaço, havendo outros elementos dividindo a mesma atmosfera); e a terceiridade, quando a mente consegue estabelecer relações entre os objetos e o ser (ALMEIDA, 2009).

As concepções peirceanas de Semiótica encontram fundamento na categorização que o referido autor estabelece para o campo da Filosofia, dividindo-a primariamente em Fenomenologia, Ciências Normativas e Metafísica. As Ciências Normativas, por sua vez,

constitui-se pela Estética, pela Ética e pela Lógica ou semiótica (que, por sua vez, se subdivide em Gramática pura ou especulativa, Lógica crítica e Metodêutica ou retórica especulativa) (SANTAELLA, 2005).

Para Ibri (2001, p. 67) a “Fenomenologia ou Faneroscopia, para C. S. Peirce, é uma ciência daquilo que aparece à experiência e, segundo o autor, trata-se da primeira das ciências da Filosofia.”

No contexto das Ciências Normativas, “voltadas para a compreensão dos fins, das normas e ideais que guiam o sentimento, a conduta e o pensamento humano” (SANTAELLA, 2005, p. 36), tem-se a Estética, a Ética e a Lógica (ou semiótica). Assim, a Estética determina o que nos é belo, tudo aquilo que é admirável e que seja “ (...) necessário abrir o espírito para aprender através das sensações e das qualidades dos objetos que surgem com a experiência” (ALMEIDA, 2009, p. 191), ao passo que, na Ética, tem-se a determinação das justificativas do que deve ser considerado certo e errado em uma sociedade, estabelecendo caminhos que devem ser seguidos por ela. Para tanto, utiliza-se da determinação do belo, ou seja, da estética (ALMEIDA, 2009).

Na Lógica ou Semiótica tem-se a teoria dos signos, este entendido como “(...) aquilo que, sob certo aspecto ou modo, representa algo para alguém. Dirige-se a alguém, isto é, cria na mente dessa pessoa, um signo equivalente, ou talvez, um signo mais desenvolvido” (PEIRCE, 2005, p. 46).

Iniciando-se pela gramática especulativa, a Semiótica, nessa etapa, tem os signos estudados enquanto forma, determinando o que devem conter para serem como tal (SANTAELLA, 2005).

Peirce ainda divide o signo em três tricotomias. A primeira tricotomia diz respeito ao relacionamento do signo com ele mesmo, podendo ser Qualissigno, o Sinsigno e o Legissigno. A segunda tricotomia refere-se ao relacionamento do signo com o seu objeto, tendo como tipos o Ícone, o Índice e o Símbolo. Já na última tricotomia, tem-se a relação do signo com seu interpretante a partir de três aspectos: Rema, Dicissigno e Argumento (PEIRCE, 2005).

Na categoria peirceana do Argumento, tem-se a lógica crítica, a partir da abdução, da dedução e da indução, em que a Abdução “formar uma predição geral sem nenhuma certeza positiva de que ela se verificará (...) como única esperança possível de regular racionalmente nossa conduta futura.” (PEIRCE, 2005, p. 60), a Dedução parte de um argumento representado como “pertencendo a uma classe geral de argumentos possíveis exatamente análogos” de tal forma que, ao longo do tempo aqueles com premissas verdadeiras levarão a

conclusões verdadeiras.” (PEIRCE, 2005, p. 59), e a Indução, como:

um método de formar Símbolos Discentes relativos a uma questão definida, cujo o método o Interpretante não representa como capaz de proporcionar, a partir de premissas verdadeiras, resultados aproximadamente verdadeiros na maioria dos casos e, no decorrer da experiência, considerando, no entanto, que se persistir nesse método, a longo prazo chega-se à verdade ou a um ponto sempre mais perto da verdade, a respeito de qualquer questão (PEIRCE, 2005, p. 60).

Desse modo, observa-se, na Abdução, o nascimento das hipóteses; na Dedução, o espaço em que as consequências lógicas são levadas em conta e, na Indução, a possibilidade de se testar a hipótese inicial.

4 A REPRESENTAÇÃO DO CONHECIMENTO E DA INFORMAÇÃO E SUA RELAÇÃO COM A SEMIÓTICA PEIRCEANA

A partir do momento que o signo representa algo, o livro, em um catálogo pode ser considerado um signo, já que ele é a representação da realidade do livro como objeto físico.

Tal processo vem sendo diferentemente denominado, na literatura da Ciência da Informação, como de criação de *substitutos de conhecimento* (CUBILLO, 199), de *substitutos documentais* (OLSON, 2002 a, b) ou de *substitutos bibliográficos* (HOWARTH, 2005), em que se supões que um documento seja representado por meio de uma série de termos, que serão interpretados por futuros usuários, sujeitos a críticas e à aceitação daquele público. Dessa forma, uma linguagem documental deve estar sempre aberta a revisões e eventuais correções, o que lembra o início das reflexões de Peirce a partir das teorias científicas e suas validações.

Resgatando a segunda tricotomia do signo - Ícone, o Índice e o Símbolo, que se refere à relação do signo com seu objeto, tem-se o índice, que sugere o objeto, mas não é o objeto em si. Observa-se, assim, uma nítida relação com a elaboração de índices pelos indexadores, uma vez que o índice sugere um conteúdo mas não é o conteúdo documental em si. Assim, o profissional indexador busca um *hábito interpretativo* em que, por meio de procedimentos mentais, procura estabelecer as diversas interpretações que podem ser dadas àqueles documentos. No entanto, ainda não se tem certo como isso é realizado e o que é nele considerado, aspectos que podem receber significativa contribuição da Semiótica (ALMEIDA e GUIMARÃES, 2008).

Os assuntos são, assim, dispostos a partir da percepção de mundo do indexador e de como eles se relacionam com determinado documento, sendo o termo indexador considerado

como signo, enquanto o assunto deve ser identificado como um interpretante e, o documento, como objeto.

Essa percepção de mundo do indexador traz consigo aquilo que Olson (2002) denomina como “poder de nomear” na medida em que o resultado da representação fica a mercê daqueles elementos que o indexador, a partir de sua visão de mundo, considera relevantes, o que caracteriza um processo eminentemente interpretativo.

Um estudo realizado por Almeida (2014), busca demonstrar através de uma coleta de dados os principais autores que trabalham com a temática e em que nível a temática é trabalhada. Pode-se perceber que autores como Peirce, Lara, Santaella e Tálamo, além de serem autores produtivos na área, são os que mais são citados. Também é importante ressaltar que, no universo pesquisado, os artigos relacionados ao tema ainda representam uma porcentagem pequena, se comparado ao número total.

No âmbito da Semiótica relacionada à representação do conhecimento, alguns autores realizaram significativas pesquisas sobre o tema, como abordado a seguir.

Representação do conhecimento

Ainda no cenário internacional, Thellefsen (2002) nos apresenta uma teoria de desenvolvimento de um modelo semiótico para organização do conhecimento, em que parte da premissa que o principal objetivo dos sistemas de organização do conhecimento é tornar o conhecimento acessível para determinada comunidade, mas, muitas vezes, esses sistemas não conseguem levar em conta as estruturas cognitivas desse conhecimento. Para o desenvolvimento científico é necessário que haja uma visão sócio-pragmática, unindo o conhecimento como produção, comunicação e interação social (THELLEFSEN, THELLEFSEN, 2004, p. 178).

Para o autor, a língua deve ser vista como independente do indivíduo em si, tendo sua evolução a partir apenas do contexto social. Assim, os significados das palavras assumem uma questão social e cultural. Dessa forma, fica mais fácil abordá-la a partir de um olhar mais cognitivo do que apenas linguístico.

Thellefsen e Thellefsen (2004) ao definirem *domínios de conhecimento*, relatam como sendo campo de conhecimento que contenha todo o conhecimento relacionado a determinado objeto ou ainda, sobre certa perspectiva. Ou seja, o conhecimento pode ser entendido como dependente de determinado ponto de vista, estabelecendo o que é significativo dentro daquela área.

Ao abordar aquilo que denomina *conceito significativo*, o autor destaca que um

conceito geralmente traz mais informações para determinadas comunidades do que para outras, o que ocorre a partir da relação entre o signo e o contexto. No âmbito científico, por exemplo, os conceitos que são referência constituem signos fundamentais, que “contém o mais vasto potencial de conhecimento e comunica o máximo de informação para o intérprete, considerando o nível de conhecimento do intérprete” (THELLEFSEN, 2002, tradução nossa, p. 77). Desse modo, representam unidades de conhecimento potenciais para aquela área, na medida em que são conceitos significativos, dos quais derivam conhecimentos posteriores, sendo que essa criação de novos conceitos só cessa, quando os conceitos começam a se tornar redundantes.

Para entender a natureza desses signos fundamentais, é necessário compreender o processo de evolução do domínio do conhecimento. Essa evolução ocorre pela constante criação de hábitos, por pessoas que possuem o mesmo objetivo dentro daquele domínio do conhecimento e para que um signo fundamental seja considerado como tal, é necessário que ele possua vários conceitos relacionados a ele, com cada um deles interpretando aspectos fundamentais do signo inicial (THELLEFSEN, 2002).

Em uma experiência na área de Terapia Ocupacional, o autor, para identificar os signos fundamentais da área e, portanto, estabelecer seu método de organização do conhecimento, o autor conversou com um grupo editorial da área para identificar qual era o signo fundamental da Terapia Ocupacional na opinião deles. Chegando ao resultado de que “atividade” poderia ser considerada o conceito fundamental, o autor perguntou a alguns professores da área se concordavam com a opinião do grupo, o que foi aceito também como opinião dos professores. A partir disso, o autor passou a estabelecer todos os conceitos relacionados à palavra atividade, para determiná-la como, de fato, um signo fundamental. Em etapa posterior, o autor isolou os conceitos relacionados à *atividade*, chegando a outros conceitos relacionados e, dessa forma, estabelecendo novos conceitos a partir de conceitos iniciais (THELLEFSEN, 2002).

Em termos semióticos, o Thellefsen (2002) propõe uma adequação da tríade peirceana para o domínio do conhecimento, onde o objeto é a manifestação do potencial do conhecimento, o fundamento como sendo o termo em si e o interpretante, o hábito interpretativo que ocorre quando se entende o significado do conceito.

O autor ainda considera os conceitos como sendo icônicos, índices e símbolos, já que eles são considerados organizadores do conhecimento, sendo imbuídos de conhecimentos. Também podem ser considerados objetos dinâmicos, pois são propensos a diferentes interpretações (THELLEFSEN, THELLEFSEN, 2004, p. 183).

A terceiridade, com seus conceitos interpretativos, é, para Thellefsen (2002), a parte mais importante para o domínio do conhecimento, uma vez que para um signo significar algo, deve haver certa regularidade subjacente e essa regularidade é proposta por meio de criação de hábitos sociais, o que é fator predominante na hora de se estabelecer um domínio de conhecimento.

Representação da Informação

Especificamente no que se refere às contribuições da semiótica peirceana para a representação do conhecimento, em especial naquilo que a corrente francesa denomina como representação documental, relativa ao conteúdo temático do documento, ressalta Lara que a abordagem linguístico-semiótica propicia uma melhor compreensão da linguagem documental como forma de organização de um conjunto de signos, assim como “suas características e forma de funcionamento enquanto sistema semiótico particular.” (LARA, 2006).

Para a referida autora, a semiose é “a passagem contínua de signo a signo, que envolve também (...) o intérprete e o contexto” sendo que “o signo é sempre uma correlação entre uma forma significante e uma unidade que definimos como significado” (LARA, 2006).

Lara e Tálamo (2007) ressaltam que a elaboração de produtos documentários é composta por uma série de atividades carregadas de valores e que, portanto, devem conter bases para que não sejam definidos aleatoriamente a partir de apenas pontos de vistas de indexadores.

As autoras ainda apontam que esse processo permite produzir uma interpretação global, onde as unidades documentárias relacionam-se com aquelas definidas para compor a linguagem documentária.

Lara (2007) destaca, ainda, que esse processo de tradução – de linguagem natural para uma linguagem controlada – envolve claramente a questão da significação e do contexto em que ela se encontra. Essa interpretação poderá ser diferenciada quando levada em conta a experiência colateral do interpretante, o conhecimento prévio que o interpretante possui sobre aquele conceito.

Para tanto, é estabelecido o conceito de signo documentário, elemento que funciona como representativo de informação, considerando o contexto em que ele se encontra (LARA, 2007). Dessa forma, propõe-se o conceito de *semiose documentária*, onde operadores semântico-semióticos localizam em determinado domínio, quais sejam: os termos e sua relação de significação. Para tanto, por meio de elementos como termos genéricos, específicos, associativos e de equivalência, registros de definições e notas de explicação,

pode-se estabelecer a interpretação do significado dos documentos recuperados (LARA, 2006).

Ainda no cenário brasileiro, Mariângela Fujita e seus colaboradores vêm desenvolvendo pesquisa também nessa área, tendo como pressuposto fundamental o fato de que “a origem sígnica do processo de indexação, sob a luz da teoria peirceana, é um recorte conceitual necessário para a ampliação de nosso horizonte investigativo” (DAL’EVEDOVE, ALMEIDA e FUJITA, 2013, p. 191), sendo que as concepções peirceanas de conceitos científico e do hábito encontram-se diretamente ligadas ao processo de indexação de assunto, uma vez que hábitos devem ser desconstruídos quando não prezarem pela recuperação documental, experiência colateral – considerando que a interpretação do indexador será alterada de acordo com as experiências colaterais vividas por aquele indivíduo, interpretação de conceitos – a tradução de assuntos para termos, e classes de signos.

Nesse âmbito, a interpretação de conceitos atua como tarefa inerente à indexação de assunto, recebendo importante subsídio da Semiótica acerca da explicação de como ocorre esse processo de interpretação, tarefa ainda pouco explorada pelos teóricos da área pois, em que pese a inexistência de um consenso na literatura da Ciência da informação sobre as etapas do processo de indexação de assunto, deve-se considerar o fato de que tais etapas visam a interpretar conceitos “de forma esquemática e operacional, além de possuírem uma interface com as classes de signos.” (DAL’EVEDOVE, ALMEIDA, FUJITA; 2013, p. 197). Tem-se, assim, um contínuo processo de interpretação de signos, que pode se estabiliza mas que nunca cessa, uma vez que as interpretações individuais irão sempre ocorrer.

Como destacam Almeida, Fujita e Reis (2013), a inferência proposta por Peirce está intimamente ligada aos processos de indexação, na medida em que os processos cognitivos para descoberta de novas informações é fundamental para estabelecer a indexação de um documento. Nesse contexto, as relações do signo consigo mesmo, do signo com o objeto dinâmico e do signo com o interpretante compõem a base da tricotomia estudada por Peirce na Gramática Especulativa, ao que se aliam os estudos de inferência desenvolvidos por Peirce no âmbito da Lógica Pura, em que a abdução constitui o estágio criativo da indexação, a dedução, o estágio geral da indexação, e a indução atua como fase de teste e continuidade da representação documental.

Desse modo, as fases da indexação de assunto podem caracterizar os tipos de signos e as tricotomias descritas por Peirce, uma vez que a indexação trabalha com a criação de novos signos e a taxonomia dos signos pode auxiliar na compreensão da natureza semiótica do processo de indexação. Nesse contexto, os indexadores são predispostos a chegar às suas

conclusões por meio de inferências, condição essencial para utilizar-se da Semiótica para conhecer a realidade.

Tais aspectos reforçam a assertiva de que a Semiótica, notadamente a partir das concepções peirceanas, é essencial no processo de indexação, uma vez que esta atividade é intelectual e, portanto, inferencial.

Em seu modelo de indexação baseado na semiótica peirceana, Jens-Erik Mai destaca o nível de interpretação subjetiva que ocorre no momento da indexação, de tal forma que não se pode falar em um modelo ou um pressuposto de como se deve indexar. Para o autor, a primeira etapa do processo de indexação consiste na análise do documento, enquanto a última etapa é a análise de assunto, diretamente ligada às necessidades dos usuários. Assim, a análise do documento baseia-se tanto no texto como no conhecimento prévio do indexador, atentando-se para as partes do texto que mais interessam aos usuários potenciais daquela unidade informacional. O autor destaca ainda que as atividades de busca dos usuários em catálogos, por meio de termos representativos, para uma futura utilização dos documentos, é em si, um ato de interpretação. (MAI, 1997, 2000, 2001).

Uma vez que as linguagens documentais, enquanto linguagens de representação de conteúdo informacional preocupam-se com a linguagem e o significado que essa linguagem pode trazer, ou seja, em como um assunto pode ser representado por meio de termos, destaca-se o fato de que a filosofia da linguagem preocupa-se com a forma em que significado é determinado e estabelecido e como essa linguagem pode representar essa realidade. Tais aspectos reforçam ainda mais a forte relação entre as preocupações da Semiótica e as preocupações da Indexação. (MAI, 2001).

Ao conceber a indexação como um conjunto de etapas ou processos interpretativos que, via de regra, se resume em: determinar os principais assuntos do documento, formular os assuntos a partir de uma linguagem natural e depois, traduzi-los para uma linguagem de indexação, o autor acrescenta, ainda, elementos importantes como a configuração do documento sob análise, os assuntos do documento tal como podem ser buscados a partir da necessidade do usuário, a organização e coesão na descrição dos assuntos, e por último, a configuração terminológica dos assuntos. No entanto, destaca a falta de tal preocupação dos guias de indexação comumente utilizados pelos bibliotecários acerca de como deve ocorrer a indexação no momento da leitura documental, ou seja, o que deve ser buscado e estabelecido como assunto, pois apenas indicam algumas fontes potenciais, aspecto que vem sendo tradicionalmente estudado por Fujita (2007, 2008, 2009).

Mai sugere a teoria de Peirce como uma importante base teórica para a indexação, por

ela incluir uma explicação de como o processo de significação ocorre, pois o foco da Semiótica não é especificamente o que o fenômeno significa mas, sim, em porque e como o significado existe (MAI, 2001). O autor lembra a distinção de Peirce entre a entrada física, como as palavras, a ideia a qual essas palavras se referem e o significado que derivam dessas palavras (MAI, 1997).

Indo além, o autor ressalta que um elemento chave na teoria de Peirce, é a ideia de Semiose Ilimitada, baseada no fato de que um signo é gerado a partir de um outro signo, produzindo um novo signo. Como novos signos sempre gerarão outros signos, esse processo pode continuar ilimitadamente (MAI, 2001). Isso, como destaca o autor, ocorre na indexação, em que novos signos são desenvolvidos a partir dos iniciais, ao que denomina como *modelo semiótico de indexação* (MAI, 1997, 2000, 2001).

O autor percebe, então, que o processo de análise de conteúdo pode ser considerado como signo, uma vez que o *representamen* é a descrição de conteúdo e o interpretante é o termo de entrada, enquanto os objetos são os recursos aos quais a análise se refere (MAI, 1997). Nesse contexto, verificam-se diferentes tipos de signos e de níveis de interpretação, uma vez que podem ser apenas ações ou, mais especificamente, podem estar relacionadas de acordo com o contexto social em que se encontram. Assim, antes de um documento ser indexado, ele foi criado dentro de um contexto, e ele próprio é resultado de diversas interpretações por parte do autor. Quando ele for indexado, deverá passar por novas interpretações pelo indexador, que deverá levar em consideração o contexto social que o documento foi criado.

Para Mai (2001), a semiose pode ser claramente encontrada dentro de três etapas de indexação: inicialmente, quando o documento torna-se objeto de análise, um objeto a ser interpretado, passando a ser um *representamen*, em que o produto é um novo signo; em uma etapa intermediária, em que o *representamen* passa a ser o conjunto de temas do documento (e não mais o próprio do documento), e a descrição de assunto passa a ser o interpretante; e finalmente, quando a descrição de assunto atua como *representamen*, tendo por *interpretante* os termos de descrição.

Resgatando as tricotomias, o autor (MAI, 2001) evidencia o fato de que podemos relacionar os elementos com as tricotomias dos signos, de tal forma que o documento constitui um argumento, os assuntos atuam como símbolo discente e a descrição de assunto por meio de termos constitui legisigno indicial discente. E justifica tais escolhas alegando que; a) o documento recebe a atribuição de argumento por ser complexo e possuir uma série de ideias, cabendo ao indexador selecioná-las para realizar uma descrição de uma maneira mais

formal; b) a escolha símbolo discente para os assuntos de um documento dá-se por conta da interpretação de um assunto ser menos complexa do que a interpretação de um documento como um todo e, c) os termos são legisigno indicial discente, uma vez que são menos complexos que o documento e a descrição de assunto, que indicam e sugerem o documento e porque dependem mais da interpretação do indexador do que dos outros dois itens.

Relativamente à atribuição de termos no que se refere às tricotomias dos signos, deve ser levado em consideração seu contexto. Assim, caso o indexador esteja familiarizado com a linguagem documental, o interpretante será uma rema e, o objeto, um índice, uma vez que será interpretado de forma fácil e aponta o objeto e as ideias associadas a ele. Porém, caso o usuário seja novo ao utilizar aquela linguagem, o interpretante seria discente enquanto o objeto seria um símbolo, considerando que o usuário necessitará interpretar de uma maneira mais complexa aquele termo, que será uma convenção dependendo do contexto cultural e social.

Para o autor, maiores estudos nessa área são necessários, principalmente dentro da análise de conteúdo documental, já que esse é o primeiro passo no momento da indexação, ressaltando o quão interpretativo e variável se apresenta todo o processo de representação de assunto, que depende, em grande medida, do contexto cultural e social do indexador (MAI, 1997,2000, 2001).

5 CONCLUSÃO

No contexto dos estudos da organização do conhecimento, os processos de representação foram tradicionalmente abordados pela área a partir de sua dimensão eminentemente prática, enquanto conjunto de fazeres, relegando a segundo plano a fundamentação teórica que poderia estar neles subjacentes, como é o caso das questões relacionadas à semiótica, assunto ainda necessita de maiores estudos.

Pode-se perceber que essa necessidade é cada vez mais iminente, uma vez que as reflexões teóricas oriundas da Semiótica auxiliam a estabelecer parâmetros para a construção de signos, enquanto representações do conhecimento, contribuindo, dessa forma, para o fortalecimento teórico metodológico da representação do conhecimento e da informação no universo da Ciência da Informação, uma vez que os processos semióticos baseados na interpretação e representação podem trazer especial aporte teórico-metodológico ao indexador, em especial no momento de identificação e de representação de assunto em documentos, deslocando-se de uma concepção até então ingênua da área no sentido de que o documento apresenta uma única e explícita tematicidade para, por outro lado, considerar

fatores mais amplos e complexos, como a experiência do indexador, suas visões de mundo e qual a melhor maneira de concatenar esses aspectos com a dimensão temática do documento a ser representado.

Demonstra-se necessário, o investimento nos teóricos da área, que se demonstraram ter a formação necessária para abordarem o tema e a maior regularidade nas publicações, uma vez que além de terem poucos teóricos que tratam sobre o assunto, ele ainda não é tratado com uma regularidade maior.

É importante se explicitar que ainda não é possível, com esse estudo, contribuir para os aspectos práticos da atuação profissional e que as discussões trabalhadas, não são as únicas abordagens sobre os temas. Faz-se necessário, portanto, uma maior abordagem sobre o assunto.

REFERÊNCIAS

- ALMEIDA, C. C. Indícios da presença de fundamentos semióticos na literatura da ciência da informação. **Scire**, v. 20, n. 1, p. 65 – 71, 2014.
- ALMEIDA, C. C.; FUJITA, M. L. S.; REIS, D. M. Peircean semiotics and subject indexing: contributions of speculative grammar and pure logic. **Knowledge Organization**, v. 39, p. 225-241, 2013.
- ALMEIDA, C. C.; GUIMARÃES, J. A. C. Análise peirceana do processo de indexação: em busca de fundamentos para a organização da informação. **In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, IX**, 2008, São Paulo. São Paulo: ENANCIB, 2008. Disponível em: <<http://www.enancib2008.com.br/>>. Acesso em 22.04.2014.
- BARITÉ, M. Organización del conocimiento: un nuevo marco teórico-conceptual en Bibliotecología y Documentación. In: CARRARA, K. (org.). **Educação, Universidade e Pesquisa**. Marília: Unesp-Marília-Publicações; São Paulo: FAPESP, 2001. p.35-60.
- DAHLBERG, I. Knowledge Organization: a new science?. **Knowledge Organization**, v. 33, n. 1, p. 11-19, 2006.
- DAL'EVEDOVE, Paula Regina ; ALMEIDA, Carlos Cândido ; FUJITA, M. S. L. . A interpretação de conceitos no processo de indexação: aspectos e reflexões semióticas. In: RIBEIRO, F., CERVEIRA, M. E.. (Org.). Informação e/ou conhecimento: as duas faces de Jano - **Atas do I Congresso ISKO Espanha e Portugal e XI Congresso ISKO España**. 1ed.Porto: Faculdade de Letras da Universidade do Porto; CETAC.MEDIA, 2013, v. 1, p. 190-203.
- FRIEDMAN, A.; THELLEFSEN, M. Concept theory and semiotics in knowledge organization. **Journal of Documentation**, v. 67, n. 4, p. 644-674, 2011.
- FUJITA, M. S. L. Abordaje cognitivo de la lectura documentaria en la formación inicial del indexador: uso del protocolo verbal en la investigación de estrategias de enseñanza. **Scire (Zaragoza)**, v. 15, p. 1-12, 2009.

FUJITA, M. S. L. La enseñanza de la lectura documentaria en el abordaje cognitivo y socio-cognitivo: orientaciones a la formación del indizador. **Anales de Documentación**, v. 10, p. 1-16, 2007.

FUJITA, M. S. L.; FERREIRA, G. I. S. Ensino do processo de análise de assunto para indexação com aplicação de um modelo de leitura: estudo de avaliação comparada em cursos de graduação em biblioteconomia no Brasil. **Ibersid (Zaragoza)**, v. 2, p. 163-176, 2008.

GUIMARÃES, J.A.C. A dimensão teórica do tratamento temático da informação e suas interlocuções com o universo científico da International Society for Knowledge Organization (ISKO). **Revista Ibero-Americana de Ciência da Informação**, v. 1, p. 77-99, 2008.

HJØRLAND, B. What is knowledge organization?. **Knowledge Organization**, v. 35, n. 3/2, p. 86-101, 2008.

IBRI, I. A. Ser e aparecer na filosofia de Peirce: o estatuto da fenomenologia. **Cognitio**, n. 2, p. 67-75, 2001.

LARA, M. L. G. Algumas contribuições da semiologia e da semiótica para a análise das linguagens documentárias. **Ciência da Informação**, v. 22, n. 3, p. 223-226, 1993a.

LARA, M. L. G. É possível falar em signo e semiose documentária?. **Encontros Bibli (UFSC)**, v. n.esp., p. 18-29, 2006.

LARA, M. L. G. Linguagens documentárias, instrumentos de mediação e comunicação. **Revista Brasileira de Biblioteconomia e Documentação**, v. 26, n. 1/2, p. 72-80, 1993b.

LARA, M. L. G. O processo de construção da informação documentária e o processo de conhecimento. **Perspectivas em Ciência da Informação**, v. 7, n. 2, p. 127-139, 2002.

LARA, M. L. G. O unicórnio (o rinoceronte, o ornitorrinco...), a análise documentária e a linguagem documentária. **DataGramaZero**, v. 2, n. 6, 2001.

LARA, M. L. G.; TALÁMO, M. F. G. M. Uma experiência na interface lingüística documentária e terminologia. **DataGramaZero**, v. 8, n. 5, 2007.

MAI, J-E. Semiotics and indexing: an analysis of the subject indexing process. **Journal of Documentation**, v. 57, n. 5, p. 591-622, 2001.

MAI, J-E. The concept of subject: on problems in indexing. In: McILWAINE, I. C. (ed.). **Knowledge organization for information retrieval: 6th International Study Conference on Classification Research**. The Hague: FID, 1997, p. 60 – 67. (FID, n. 716).

MAI, J-E. **The subject indexing process: an investigation of problems in knowledge representation**. 2000. 344 f. Tese (Doutor em Filosofia) - Faculty of Graduate School of Library and Information Science, The University of Texas at Austin.

NOTH, W. **Semiótica do século XX**. São Paulo: Annablume, 1996.

PEIRCE, C. S. **Semiótica**. Tradução por José Teixeira Coelho Neto. São Paulo: Perspectiva, 2005.

PEIRCE, C. S. **The essential Peirce: selected philosophical writings**. Indiana University Press, 1998.

SALES, R. et. al. Redes sociais em linguagens documentais: uma análise de coautoria a partir da realidade brasileira. **Encontros Bibli**, v. 16, n. 31, p. 1-24, 2011.

SANTAELLA, L. Matrizes da linguagem e pensamento: sonora, visual, verbal: aplicações na hipermídia. 3. Ed. São Paulo: Iluminuras, FAPESP, 2005.

TÁLAMO, M. F. G. M. Knowledge production, interdisciplinarity, and structuralism. **Electronic Journal of Communication Information & Innovation in Health**, v. 3, n. 3, p. 119-126, 2009.

TÁLAMO, M. F. G. M. Terminologia e documentação. **TradTerm**, v. 7, p. 141-151, 2001.

TÁLAMO, M. F. G. M.; LARA, M. L. G. M. Interfaces between linguistics, terminology and documentation. **Brazilian Journal of Information Science**, v. 3, n. 2, p. 54-70, 2009.

TÁLAMO, M. F. G. M.; LARA, M. L. G. M. O campo da lingüística documentária. **TransInformação**, v. 18, n. 3, p. 203-211, 2006.

THELLEFSEN T. L.; THELLEFSEN, M. M. Pragmatic semiotics and knowledge organization. **Knowledge Organization**, v. 31, n. 3, p.177 – 187, 2004.

THELLEFSEN, T. Semiotic knowledge organization: theory and method development. **Semiotica**, v. 142, n. 1/4, p. 71-90, 2002.

PRINCÍPIOS METODOLÓGICOS PARA DESENVOLVIMENTO DE ONTOLOGIAS: ANÁLISE DAS PRÁTICAS CORRENTES E PROPOSIÇÃO DE MELHORIAS

METHODOLOGICAL PRINCIPLES FOR BUILDING OF ONTOLOGIES: ANALYSIS OF THE CURRENT PRACTICES AND SUGGESTION OF IMPROVEMENTS

Fabício Martins Mendonça
Maurício Barcellos Almeida

Resumo: O montante de pesquisas e aplicações práticas sobre o tema ontologia continua crescendo no âmbito da Ciência da Informação e em áreas afins, principalmente devido as novas possibilidades que ontologias proporcionam na representação e organização do conhecimento. Uma grande aplicação mundial de ontologias tem sido a estruturação de conhecimento no domínio biomédico, por exemplo, com o surgimento e crescimento de repositórios para integração de ontologias biomédicas. Essa efervescência da atividade de desenvolvimento de ontologias pode ser notada também no número de metodologias criadas para esse fim. Entretanto, ainda não existe uma metodologia unificada que seja amplamente aceita, uma vez que alguns problemas continuam em aberto mesmo nas principais metodologias em uso. Diante deste cenário, o presente artigo apresenta uma análise descritiva das metodologias atuais, evidenciando aspectos ontológicos básicos que são essenciais à uma metodologia genérica de desenvolvimento de ontologias. Como resultado dessa análise, propõe-se um novo conjunto de princípios metodológicos, em teste no domínio biomédico relacionado ao sangue humano, o qual leve em conta em seus procedimentos os aspectos ontológicos básicos evidenciados.

Palavras-chave: Ontologia. Metodologia de desenvolvimento de ontologias. Representação do Conhecimento.

Abstract: The amount of research and of practical applications involving the theme ontology has been growing in the scope of Information Science, as well as in relates scientific fields, mainly due the new possibilities that ontologies provide in knowledge organization and representation. A big world application of ontologies has been the structuration of knowledge in the biomedical domain, for example, with the emergence and growth of the repositories for integration of biomedical ontologies. This effervescence of the activity of ontologies development might be noticed in the number of methodologies created for this purpose. However, there is no unified methodology broadly accepted, insofar as some issues are still open, even considering the main current methodologies. In this context, this article presents a descriptive analysis of the current methodologies for ontology development, emphasizing some basic ontological aspects that are essential to a general methodology for ontologies development. As a result of that analysis, we propose a new set of methodological principles, in test in the biomedical domain related to human blood, which take into account those ontological aspects emphasized by the previous analysis.

Keywords: Ontology. Methodology for development of ontologies. Knowledge Representation.

1 INTRODUÇÃO

O estudo de ontologias caracteriza-se como um ramo de pesquisa que surgiu no final dos anos 80, propondo alternativas para representar o conhecimento. Usadas como instrumentos para a modelização do conhecimento de domínio diversos, as ontologias

apresentam uma natural proximidade com técnicas consagradas pela Ciência da Informação. Os primeiros trabalhos sobre essa temática na área são identificados em Soergel (1997) e Vickery (1997).

Desde então, o interesse e a aplicação prática de ontologias em pesquisas inseridas no contexto da Ciência da Informação vem aumentando consideravelmente e seu uso já está difundindo em diversos ramos do conhecimento, que incluem, por exemplo: (i) suporte à extração de informação de grandes bibliotecas disponíveis em fontes diversas da web; (ii) tradução da linguagem natural; (iii) integração automática de um conjunto de vocabulários padronizados ou dicionários de dados relacionados a um domínio específico; (iv) integração de bancos de dados, de softwares ou de modelos de negócio; (v) recuperação de informação médica e processamento dos registros eletrônicos dos pacientes; (vi) aplicações práticas em recursos da web semântica; entre outros (SMITH, 2004).

Esse mesmo autor afirma ainda que o aumento do uso de ontologias é mais notável na área biomédica, onde desempenham um papel fundamental como iniciativa de padronização do vocabulário médico. Embora essa função seja semelhante àquela das terminologias médicas, as ontologias biomédicas são instrumentos de representação mais expressivos que as terminologias considerando o uso por computadores, uma vez que utilizam formalismos baseados em lógica. Além disso, possibilitam a representação de um domínio do conhecimento mais próxima da realidade, o que é essencial no caso da biomedicina.

Na verdade, a maior aplicação mundial de ontologias tem sido a estruturação de conhecimento no domínio biomédico. A expansão na construção e utilização de ontologias biomédicas pode ser notada pelo surgimento de repositórios especializados, tal como o *Open Biomedical Ontologies* (OBO) (SMITH et al., 2007), que visam à integração e padronização de ontologias biomédicas, disponíveis *on line* em domínio público.

Mesmo com esse crescimento notável nas atividades relacionadas às ontologias, alguns problemas no desenvolvimento de ontologias permanecem em aberto, talvez em função da imaturidade de uma área de pesquisa recente. Muitos desses problemas podem ser observados nas metodologias atualmente disponíveis para a construção de ontologias. De forma geral, não há um padrão de construção ou uma metodologia unificada que seja amplamente aceita para o desenvolvimento de ontologias. Tais metodologias, ainda hoje, não são consideradas suficientemente “maduras” como as metodologias empregadas na engenharia de software ou na construção de tesouros, por exemplo (USCHOLD e GRUNINGER (1996); JONES, BENCH-CAPON, VISSER (1998); FERNANDÉZ-LOPEZ

(1999); FERNANDÉZ-LOPEZ e CORCHO (2004); BREITMAN (2005); CARDOSO (2007); SILVA, SOUZA, ALMEIDA (2008)).

Dada essa questão, esse artigo apresenta uma análise das principais metodologias de construção de ontologias em uso, evidenciando aspectos ontológicos essenciais que, de alguma forma, deveriam ser tratados e incluídos em uma metodologia genérica para a construção de ontologias. Como resultado dessa análise, o artigo propõe um novo conjunto de princípios metodológicos para a construção de ontologias, especialmente, em domínios biomédicos, que leve em consideração os aspectos ontológicos abordados.

Enfim, a ideia do conjunto de princípios proposto é unificar as etapas de algumas das metodologias atuais, tais como a *Methontology* (GÓMEZ-PEREZ, FERNANDEZ-LOPES e VICENTE, 1996), o ciclo de vida de *Schiessl e Bräscher* (2011) e o método 101 (NOY e GUINNESS, 2001), fazendo adaptações necessárias aos domínios biomédicos e acrescentando detalhes em cada uma das etapas.

O restante do presente artigo está assim organizado: a seção dois apresenta e descreve, brevemente, as principais metodologias correntes para o desenvolvimento de ontologias, bem como analisa o cenário de problemas encontrado. A seção três evidencia aspectos essenciais, baseados em princípios ontológicos, que deveriam fazer parte de qualquer metodologia genérica de desenvolvimento de ontologias. A seção quatro apresenta o conjunto de princípios metodológicos proposto, organizado na forma de uma nova metodologia. Finalmente, a seção cinco apresenta considerações finais e sugestões de trabalhos futuros.

2 METODOLOGIAS PARA CONSTRUÇÃO DE ONTOLOGIAS

Na literatura da área ontológica é possível identificar uma série de metodologias e métodos para a construção de ontologias que, em geral, são direcionadas para aplicações em domínios específicos. A revisão de literatura e o estudo das metodologias e métodos mais utilizados atualmente na área de engenharia ontológica foram realizados, na presente pesquisa, com o propósito de apreender o background teórico necessário para a construção de uma nova metodologia para a construção de ontologias.

Os objetivos desta seção do artigo são: apresentar, de maneira resumida, a revisão de literatura das metodologias e métodos pesquisados e, em seguida, realizar uma análise de tais metodologias e métodos, tendo como base seus aspectos e fundamentos principais, além de considerar nessa análise as avaliações de tais instrumentos já realizadas por autores da área.

2.1 Descrição das metodologias atuais

Através da literatura da área, é possível perceber que houve uma proliferação de metodologias e métodos de construção entre meados da década de 90 até o início dos anos 2000. Nesse período surgiram as metodologias até hoje mais referenciadas por pesquisas na área, tais como: a metodologia de Gruninger e Fox – TOVE (GRUNINGER e FOX, 1995), a *Methontology* (GÓMEZ-PEREZ, FERNANDEZ-LOPES, VICENTE, 1996), o método 101 (NOY e GUINNESS, 2001), entre outras. De meados dos anos 2000 até hoje, surgiram novas metodologias para construção de ontologias, tais como: a metodologia *UP for ONtology* (UPON) (DE NICOLA, MISSIKOFF, NAVIGLI, 2009) e a metodologia *NeON* (SUARÉZ-FIGUEROA, 2010), entretanto, em um número menor se comparado ao período anterior.

Mesmo que tais metodologias apresentem algumas semelhanças entre seus aspectos principais e etapas previstas para a construção de ontologias, podem ser percebidas diferenças entre elas, principalmente, quando é realizada a comparação entre metodologias direcionadas a domínios de aplicação distintos entre si.

A fim de compreender melhor essas metodologias e métodos, construiu-se uma tabela descritiva (QUADRO 1) sobre tais instrumentos, destacando como seus aspectos principais as seguintes informações: (i) uma breve descrição da metodologia, por vezes, incluindo seu objetivo principal; (ii) o domínio de aplicação ao qual são indicadas, que está incluído na descrição; e, de forma resumida, (iii) as etapas, que incluem processos e atividades, de cada uma das metodologias pesquisadas. Essas informações estão apresentadas no QUADRO 1, a seguir, o qual apresenta tais metodologias e métodos em ordem cronológica de sua criação. Para maiores detalhes sobre as etapas de cada metodologia, sugere-se consultar as referências fornecidas.

QUADRO 1 - Metodologias de construção de ontologias

Metodologia	Descrição (objetivo e domínio)	Etapas (processos e atividades)
Metodologia de Gruninger e Fox – TOVE (GRUNINGER e FOX, 1995)	Criada no projeto <i>Toronto Virtual Enterprise (TOVE)</i> , cujo objetivo era criar um modelo de senso comum compartilhado corporativo. Domínio de aplicação: Negócios (empresarial)	1. Elaboração de cenários de motivação 2. Questões de competência informal 3. Concepção da terminologia formal 4. Questões de competência formal 5. Especificação de axiomas formais; 6. Verificação de teoremas completos.
Metodologia de Uschold e King ENTERPRISE (USCHOLD e KING, 1995)	Desenvolvido com base na prática da construção da ontologia <i>Enterprise Ontology</i> , que descreve conhecimento no domínio corporativos. Domínio de aplicação: Negócios (empresarial)	1. Propósito, grau de formalismo e usuário 2. Construção da ontologia em três etapas: a) conceitualização b) implementação c) integração com ontologias já existentes; 3. Avaliação 4. Documentação

<i>Methontology</i> (GÓMEZ-PEREZ, FERNANDEZ-LOPES e VICENTE, 1996)	Possibilita a construção de ontologias por reengenharia, utilizando-se do conhecimento de domínio. Domínio de aplicação: Diversos.	1. Especificação; 2. Aquisição de Conhecimento 3. Conceitualização 4. Integração 5. Implementação 6. Avaliação 7. Documentação
Método <i>Kactus</i> (BERNARAS, LARESGOTTI, CORERA, 1996)	Método derivado do projeto Kactus que permite reutilização de conhecimento em sistemas técnicos, tal como redes elétricas. Domínio de aplicação: Sistemas de complexidade técnica.	1. Lista de necessidades ou requisitos 2. Identificação de termos relevantes 3. Criação de modelo preliminar 4. Estruturação e refinamento 5. Integração e reutilização.
Método <i>Sensus</i> (SWARTOUT et al., 1996).	Método derivado da ontologia <i>Sensus</i> , desenvolvida para de processamento de linguagem natural. Domínio de aplicação: Diversos.	Identificar termos-chave do domínio; Ligar os termos-chave à ontologia <i>Sensus</i>; Adicionar novos termos para o domínio; Adicionar subárvores completas.
Método 101 (NOY e GUINNESS, 2001)	Método usado no desenvolvimento de exemplos práticos utilizando o editor de ontologias Protégé. Domínio de aplicação: Diversos.	1. Determinar escopo 2. Considerar o reuso de termos 3. Enumerar termos 4. Definir classes 5. Organizar as classes em uma taxonomia 6. Definir propriedades e suas restrições 7. Adicionar valores de instâncias
Método CYC (REED, LENAT, 2002)	Método usado na construção da ontologia CYC, que busca abranger o considera o conhecimento consensual do mundo. Domínio de aplicação: Diversos.	1. Extração manual do conhecimento 2. Extração auxiliada por computador 3. Desenvolvimento de representação 4. Representação do conhecimento de diferentes domínios usando primitivas.
<i>On-to-Knowledge Methodology (OTKM)</i> (SURE, STAAB e STUBER, 2003)	Metodologia desenvolvida para a construção de ontologias para aplicações de gestão do conhecimento empresarial. Domínio de aplicação: Gestão do conhecimento empresarial.	1. Estudo de viabilidade 2. <i>Kickoff</i> 3. Refinamento 4. Avaliação 5. Aplicação e evolução 6. Geração de conhecimento 7. Captura de conhecimento 8. Recuperação e acesso
Metodologia <i>UP for ONtology</i> (UPON) (DE NICOLA, MISSIKOFF e NAVIGLI, 2009)	Metodologia de construção de ontologias derivada em padrão de engenharia de software Processo Unificado Domínio de aplicação: Negócios (<i>e-bussiness</i>).	1. Workflow de Requisitos 2. Workflow de Análise 3. Workflow de Desenvolvimento 4. Workflow de Implementação 5. Workflow de Teste
Metodologia <i>NeON</i> (SUARÉZ-FIGUEROA, 2010)	Metodologia para construção de redes ontológicas baseado em um desenvolvimento colaborativo Domínio de aplicação: Diversos.	1. Gerenciamento de processos e atividades 2. Desenvolvimento orientado de processos e atividades 3. Suporte aos processos e atividades

<p>Metodologia <i>MFPFO</i> (LIM, LIU e LEE, 2011)</p>	<p>Metodologia de construção de ontologia multi-facetada, anotada semanticamente, para a modelagem de uma família de produtos. Domínio de Aplicação: Domínios que possuem uma família de produtos</p>	<ol style="list-style-type: none"> 1. Construção de uma taxonomia da família de produtos; 2. Extração de entidades; 3. Identificação do conceito e geração da unidade facetada; 4. Anotação semântica e modelagem faceta; 5. Construção de uma ontologia de família de produtos multi-facetada e anotada semanticamente; 6. Avaliação e validação da ontologia.
<p>Ciclo de Vida de Schiessl e Bräscher (2011)</p>	<p>Embora não seja uma metodologia, o ciclo de vida ontológico inclui todas as etapas necessárias no processo de construção de ontologias Domínio de Aplicação: Diversos.</p>	<ol style="list-style-type: none"> 1. Especificação 2. Conceitualização 3. Formalização 4. Aplicação 5. Manutenção 6. Atividades paralelas: aquisição de conhecimento, avaliação, documentação

Fonte: elaborado pelos autores.

2.2 Análise do cenário atual das metodologias

Após a revisão realizada sobre as metodologias para construção de ontologias, procedeu-se com uma análise do cenário atual das metodologias, destacando problemas, identificados tanto na literatura da área quanto na análise realizada a partir de prática, e também considerações sobre o uso de tais metodologias em pesquisas.

Uma consulta às etapas das metodologias pesquisadas (QUADRO 1) permite verificar que existem diferenças entre elas, sendo algumas mais completas e detalhadas – tais como a *Methontology* e a *TOVE* – e outras mais simples, que dão ênfase a aspectos de implementação da ontologia – tais como os métodos 101, *CYC* e *Kactus*.

Santos (2014) afirma que três metodologias têm se destacado mais na construção de ontologias de domínio: o método 101, a *On-to-knowledge* e a *Methontology*. Destaca ainda que as metodologias de construção atuais não são usadas isoladamente, mas em parceria com outras metodologias como forma de suprir as limitações existentes em cada uma delas, como por exemplo, ocorre com as metodologias *Methontology*, *On-to-knowledge* e o método 101. Dessa forma, metodologias próprias estão ganhando espaço na literatura da área, uma vez que os desenvolvedores e pesquisadores estão fazendo adaptações e elaborando novas metodologias, tal como é realizado na criação da metodologia de construção proposta nesta pesquisa.

A elaboração de uma metodologia própria para construção de ontologias é justificada também pelos problemas identificados nas metodologias atuais. Alguns dos principais problemas que podem ser relatados são:

- I. Em relação a detalhes das atividades, algumas **metodologias e métodos mostram-se superficiais na descrição dos passos para construção de ontologias**, considerando que o ontologista já domina o assunto sobre e não necessita de detalhes acerca de atividades e procedimentos. Um exemplo disso é a ausência de orientações que ajudem em decisões ontológicas sobre a inclusão ou não de conhecimento na ontologia (SILVA; SOUZA; ALMEIDA, 2008).
- II. **Algumas abordagens dão mais ênfase em atividades de desenvolvimento**, especialmente a implementação da ontologia (método Cyc e método 101), desconsiderando aspectos importantes relacionados a gerenciamento do projeto, a estudo de viabilidade, à manutenção e à avaliação de ontologias (SILVA; SOUZA; ALMEIDA, 2008).
- III. Outra questão em aberto sobre as metodologias de construção de ontologias diz respeito ao seu **grau de maturidade**. Para alguns autores, como, por exemplo, Fernández-López (1999), o grau de maturidade de uma metodologia para ontologias deveria ser medido a partir de sua similaridade com os padrões de desenvolvimento de software. O problema dessa abordagem é que ela considera uma ontologia apenas como componente de software. Nesse sentido, essa abordagem só funcionaria para ontologias computacionais e não para toda a gama de ontologias existentes.
- IV. **Muitas das metodologias atuais assumem uma tarefa como ponto de partida na construção da ontologia**, por exemplo, a forma de aquisição de conhecimento ou avaliação a partir de questões de competência. Se por um lado essa estratégia é interessante, por outro ela não pode vir a ser quando ela impõe limitações na reutilização da ontologia (JONES; BENCH-CAPON; VISSER, 1998).
- V. **As atividades necessárias para reutilização e integração de ontologias**, por exemplo, em um repositório ontológico, **são ainda tarefas não triviais e para as quais ainda não existe um padrão amplamente aceito para realizá-las**. Jones, Bench-Capon e Visser (1998) destacam a importância de se selecionar uma ontologia adequada e de como estender a ontologia nesse processo de reutilização.

3 ASPECTOS ESSENCIAIS NA CONSTRUÇÃO DE ONTOLOGIAS

Além da análise e revisão de literatura sobre as atuais metodologias de construção, outro tópico discutido no presente artigo e levado em consideração na criação de uma nova metodologia de construção refere-se aos aspectos ontológicos considerados essenciais no processo de desenvolvimento de ontologias.

Embora na literatura da área não seja possível identificar um número significativo e relevante de pesquisas sobre esse tema específico, fazendo um levantamento bibliográfico sobre pesquisas que abordam o tema, direta ou indiretamente, foi possível enumerar tais aspectos a partir da observação, na prática de construção de ontologia no domínio biomédico – a HEMONTO (MENDONÇA; ALMEIDA, 2013) –, dos principais problemas e omissões das atuais metodologias.

O primeiro aspecto essencial refere-se a uma **delimitação clara e explícita do escopo** de cobertura da ontologia. Esse aspecto é levado em consideração na maioria das metodologias de construção atuais. Metodologias como a *Methontology*, Gruninger e Fox, Uschold e King e método 101, citam a definição do escopo da ontologia como uma etapa de especificação, na qual são identificados os propósitos, usuários, graus de formalismo e aplicações da ontologia, além de recomendar o uso de questões de competência na definição do escopo. É necessário complementar essas atividades com uma delimitação clara e explícita do escopo. Delimitação clara e explícita significa especificar formalmente o ponto de partida e o limite do domínio cobertos, incluindo os níveis de conhecimento que serão abordados.

O segundo aspecto sugerido é a avaliação da possibilidade de **uso de ontologias de fundamentação** na construção da ontologia pretendida. Como as ontologias de fundamentação (ou ontologias de alto nível) descrevem conceitos gerais independentes do problema ou domínio (GUARINO, 1998), elas têm sido usadas como ponto de partida para a construção de ontologias sobre domínios específicos. Esse tipo de abordagem é uma prática de pesquisa incentivada na literatura sobre ontologias, especialmente as biomédicas (SMITH, 2005; BITTNER; DONNELLY, 2007; DEGEN *et al.*, 2001; MASOLO *et al.*, 2003; GRENON; SMITH, 2004; HERRE, 2006). As metodologias para construção de ontologias pesquisadas não incluem explicitamente como uma de suas etapas a avaliação de ontologias de fundamentação como ponto de partida para a construção da ontologia requerida. Entretanto, acredita-se que tal atividade seja fundamental para o desenvolvimento de ontologias.

O terceiro aspecto essencial na construção ontológica refere-se ao **reuso de termos de outras ontologias** existentes e de que forma se dá esse reuso na ontologia em desenvolvimento. Esse aspecto é contemplado pela maioria das metodologias atuais, mas a questão nesse reuso é a forma pelo qual ele é conduzido.

Uma das tarefas mais difíceis para o desenvolvedor da ontologia é decidir sobre a escolha do termo a ser reaproveitado, em função do grande número de ontologias disponíveis,

de sinônimos e de ambiguidades. Para auxiliar nessa tarefa, existem ferramentas de busca em ontologias, tais como: *Foundational Model Explorer*²⁷, Portal do *National Center for Biomedical Ontology*²⁸ (NCBO), *Ontology Lookup Service*²⁹ (OLS), *Ontobee*³⁰, que possibilitam buscar o termo pesquisado em repositórios de ontologias e obter propriedades acerca do termo. Após a seleção do termo, pode-se utilizar um editor de ontologias, tal como o *Protégé*³¹, para importação do termo para a ontologia em desenvolvimento.

De maneira geral, esses detalhes no reaproveitamento do uso de termos para ontologia não são abordados nas metodologias de construção pesquisadas. Desta forma, a metodologia proposta nesta pesquisa visa auxiliar nesse processo de reaproveitamento.

O quarto aspecto essencial trata-se da existência de uma **descrição informal e outra formal da ontologia** em desenvolvimento. Entende-se como descrição informal da ontologia a representação textual e gráfica da conceitualização do domínio, que inclui, por exemplo, o dicionário de classes da ontologia, as definições textuais de cada classe, as taxonomias das classes, entre outros. Já a descrição formal da ontologia corresponde à representação do conteúdo ontológico em uma linguagem formal, tais como a lógica descritiva, a lógica de primeira-ordem, a lógica da linguagem *Ontology Web Language* (OWL), que fornecem recursos para a especificação de axiomas, propriedades e questões de competência da ontologia de maneira formal.

Esse quarto aspecto está incluído em quase todas metodologias atuais de construção de ontologias e, muitas vezes, referenciado como a etapa de implementação da ontologia. Ele é fundamental para assegurar que a ontologia englobe o conhecimento do domínio e para tornar possível a interpretação por máquinas via inferências. Para tanto, um aspecto primordial nessas descrições formal e informal da ontologia é distinguir qual parte do conhecimento do domínio pode ser representável formalmente, e qual parte do conhecimento deve ser representado em anotações e comentários. Embora essa distinção seja fundamental no processo de construção de ontologias, a maioria das metodologias existentes apenas prescreve a existência dos dois tipos de descrição, não explicando como o conhecimento do domínio pode ser representado ou não em cada um desses tipos.

²⁷ Disponível em: <http://fme.biostr.washington.edu/FME/>. Acesso em: 08 de Junho de 2014.

²⁸ Disponível em: <http://bioportal.bioontology.org/>. Acesso em: 08 de Junho de 2014.

²⁹ Disponível em: <http://www.ebi.ac.uk/ontology-lookup/>. Acesso em: 08 de Junho de 2014.

³⁰ Disponível em: <http://www.ontobee.org/>. Acesso em: 08 de Junho de 2014.

³¹ Disponível em: <http://protege.stanford.edu/>. Acesso em: 2 jul. 2014.

O quinto aspecto considerado essencial corresponde à **criteriosidade na transição da descrição informal da ontologia para a descrição formal**. O que se quer dizer com “criteriosidade” nesse caso é tornar totalmente equivalentes, a medida do possível, a descrição formal e a descrição informal, uma vez que a primeira é derivada da segunda. Pela análise feita das metodologias, nenhuma delas deixa claro como deve ser feita essa transição entre tais descrições.

Um exemplo claro de como ser criterioso nessa transição entre os dois tipos de descrição de uma ontologia está na **elaboração das definições textuais** (em linguagem natural) das classes de uma ontologia. Num primeiro momento pode parecer trivial criar uma definição em linguagem natural para um termo (classe) da ontologia, uma vez que o senso comum sugere que uma definição em linguagem natural é apenas a descrição do significado do termo. Entretanto, para ser criterioso na transição entre a descrição informal e a formal, as definições textuais das classes devem ser elaboradas de forma a facilitar as definições posteriores em lógica. O uso de algum método para definir um termo, por exemplo, o método aristotélico, citado em Smith (2013), facilita a transição para definições formais. O uso deste método relembra, por exemplo, a considerar na definição da classe apenas as condições necessárias e suficientes que a compõem.

A tarefa de elaboração das definições textuais e formais talvez seja a mais complexa atividade na transição entre a descrição informal e a formal. Mas existem outros fatores importantes, tais como os **princípios da mono-hierarquia e da não circularidade**. O princípio da mono-hierarquia diz que cada classe da ontologia deve ter exatamente uma única classe superior, não sendo permitida a herança múltipla; o princípio da não circularidade diz que não se deve usar uma classe na definição da mesma classe (SCHULZ *et al.*, 2009; SMITH, 2013).

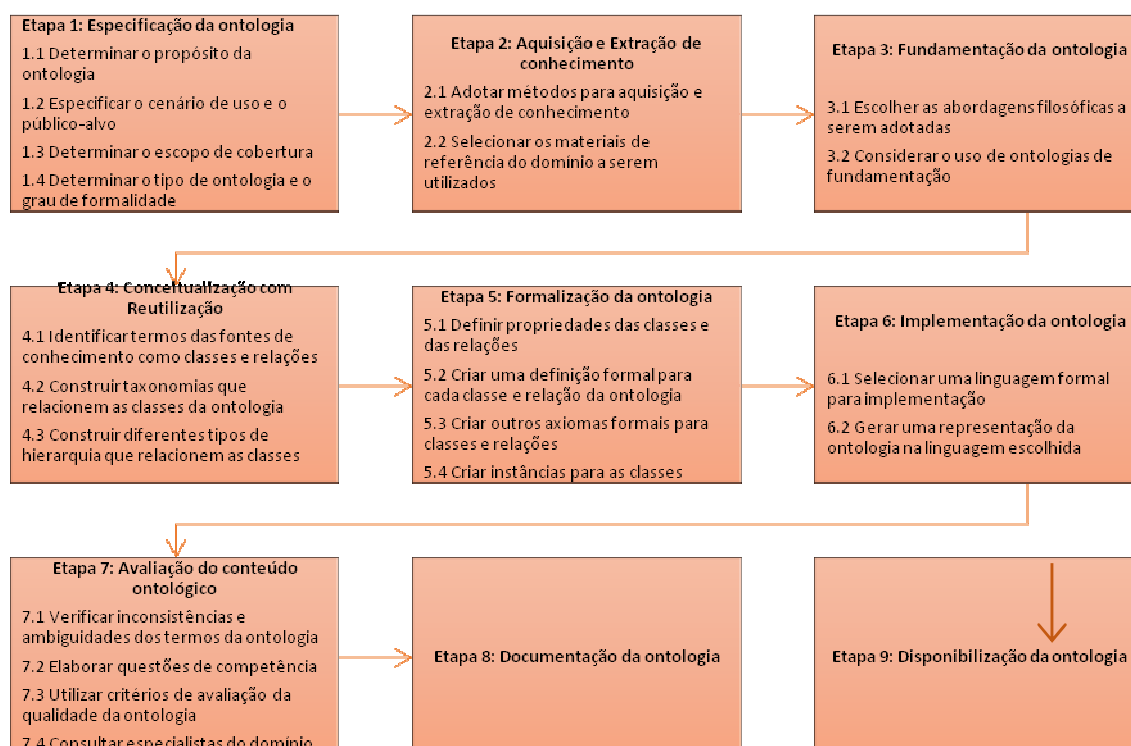
Por fim, sugere-se como sexto aspecto essencial a **completude mínima necessária do conteúdo ontológico** de cada classe e cada relação da ontologia. Nesse caso, o que se quer dizer com a expressão “completude mínima necessária” de cada classe e relação da ontologia é que a descrição de cada termo na ontologia deve englobar as informações mínimas e suficientes para representar na ontologia o objeto real a que ele se refere. Esse último aspecto sugerido, normalmente, não é tratado de forma explícita nas metodologias de construção pesquisadas, que, de uma maneira e outra, descrevem algumas informações que devem estar presentes nas classes (metadados).

4 PRINCÍPIOS METODOLÓGICOS PROPOSTOS

Após a revisão e a análise das atuais metodologias para construção de ontologias e o levantamento de alguns aspectos ontológicos essenciais no desenvolvimento de ontologias tornou-se possível criar uma fundamentação própria que serve de base para a elaboração de uma metodologia própria para construção de ontologias. Este tópico do presente artigo concentra-se em apresentar tal metodologia proposta, explicando cada uma das etapas nela incluída, e representa o resultado teórico desta pesquisa.

A fim de atingir os propósitos pretendidos com a metodologia proposta, foram reaproveitados alguns passos fundamentais no processo de construção de ontologias presentes em algumas das atuais metodologias de construção, especificamente a *Methontology* (GÓMEZ-PEREZ; FERNANDEZ-LOPES; VICENTE, 1996), a metodologia de Gruninger e Fox – TOVE (GRUNINGER; FOX, 1995), o método 101 (NOY; GUINNESS, 2001) e o ciclo de vida de Schiessl e Bräscher (2011). Nesse sentido, segue-se as etapas básicas do processo de construção incluídos em tais ontologias, porém detalhando-as com sua subdivisão em passos menores (subetapas) e adaptando-as ao domínio tratado, de forma a considerar os aspectos essenciais levantados, muitas vezes omitido nessas metodologias. As etapas incluídas na metodologia proposta estão elucidadas no diagrama da FIG. 1.

FIGURA 1 - Etapas da metodologia própria de construção de ontologias.



Fonte: elaborado pelos autores.

Na **primeira etapa** deve-se definir de forma clara e concisa o propósito geral da ontologia e seus objetivos específicos, o tipo da ontologia desenvolvida (ontologia de domínio ou de fundamentação, ontologia leve ou pesada, ontologia de sistemas de informação ou ontologia para sistemas de informação (GUARINO, 1998; GÓMEZ-PÉREZ, 2004; FONSECA, 2007)), seu escopo de cobertura, cenários de uso ou aplicação da ontologia, para que tipos de usuários ela se destina e o grau de formalidade.

Na **segunda etapa** sugere-se a adoção de métodos de aquisição de conhecimento para a construção de ontologias, tais como: (i) análises de textos em materiais de referência do domínio; (ii) entrevistas estruturadas ou semi-estruturadas com especialistas da área; (iii) grupos focais com especialistas da área; (iv) aplicações de questionários; (v) análise e técnicas de indução; (vi) *brainstorming* com técnicos e especialistas da área. Além da adoção de tais métodos é fundamental também fazer uma seleção apropriada dos materiais de referência a serem utilizados para obtenção do conhecimento do domínio tratado, que, na prática, traduz-se na escolha dos materiais de referência mais relevantes e representativos acerca do domínio, normalmente, contando-se com a ajuda de especialistas nele.

A etapa de fundamentação da ontologia (**terceira etapa**) envolve a escolha de uma ou mais abordagens filosóficas a serem adotadas no processo de desenvolvimento da ontologia e a avaliação do uso e aplicação de uma ou mais ontologias de fundamentação no desenvolvimento da ontologia pretendida. Embora essa etapa possa ser questionável por parte de ontologistas que desenvolvam ontologias essencialmente computacionais, ela é imprescindível como forma de tornar o mais transparente possível as escolhas ontológicas tomadas no processo de construção da ontologia.

A **quarta etapa** corresponde a chamada conceitualização com reutilização, cujo propósito é o de descrever em modelo conceitual (conjunto de termos e seus relacionamentos no domínio) a ontologia a ser construída de forma a atender às especificações das etapas anteriores, dando ênfase à reutilização de termos de outras ontologias (por exemplo, as ontologias de fundamentação) na ontologia em desenvolvimento. O reaproveitamento de termos (classes e relações) é um dos princípios universalmente aceitos na construção de ontologias. Como atividades dessa etapa têm-se: (i) a identificação de termos do domínio para a ontologia; (ii) a construção de taxonomias relacionando classes do domínio; e (iii) a construção de outras hierarquias (partonomias e estruturas gerais de representação) relacionando tais classes. Sobre a identificação de termos é importante destacar que ela deve ser conduzida considerando o conceito de reuso, sendo assim, sugere-se os seguintes passos nessa identificação: (a) identificar termos relevantes do domínio nos materiais de referência

utilizados; (b) pesquisar em ontologias existentes por um termo igual ou sinônimo ao termo identificado; (c) se o termo não foi encontrado então criar uma nova classe específica da ontologia, porém se o termo foi encontrado então reaproveitá-lo da ontologia de origem. Sobre as **taxonomias** é importante ressaltar que tais estruturas devem conter classes relacionadas unicamente através da relação ontológica *is_a* (*é_um*), de maneira que se defina formalmente o tipo de relação *is_a* utilizado na construção das taxonomias, uma vez que existe mais de um tipo dessa relação. A mesma regra se aplica para as **partonomias**, que devem conter classes relacionadas unicamente através da relação ontológica *todo_parte*, que também deve ter seu tipo especificado formalmente. Por fim, o que se convencionou chamar de **estruturas gerais** de representação são estruturas que englobam classes relacionadas através de diferentes relações ontológicas, como, por exemplo, as relações *produces*, *participates_in*, *has_agent* da ontologia *Relation Ontology* (RO) e as próprias relações *is_a* e *part_of* das taxonomias e partonomias, respectivamente.

A **quinta etapa** da metodologia proposta é a etapa de formalização da ontologia modelada nos passos anteriores. Tal formalização corresponde a uma descrição formal da ontologia em linguagem formal capaz de representar a conceitualização do domínio modelada na etapa anterior, ainda que essa linguagem não seja a forma final a ser usada na implementação da ontologia. Com a evolução dos editores de ontologia, tal como o Protégé 4.3, essa etapa pode ser praticamente toda realizada com uso desses softwares e envolve as seguintes atividades: (i) definição formal das propriedades das classes e das relações: para a completude mínima dessas propriedades, sugere-se o preenchimento das seguintes informações (atributos): (a) identificador da classe (*ID*); (b) nome da classe (*label*); (c) ontologia de origem (*imported from*; *hasOBONamespace*); (d) URI (*hasURI*); (e) sinônimos do termo (*hasSynonym*, *hasExactSynonym*, *hasRelatedSynonym*); (f) termo alternativo (*alternative term*); (g) definição textual (em linguagem natural) (*definition*); (h) elucidação (*elucidaton*); (i) exemplo de uso (*example of usage*); (j) definição formal (em linguagem lógica) (*SubClass Of*); (l) anotações ou comentários (*comments*). Tais atributos são originados de ontologias como *Gene Ontology* (GO), *Information Artifact Ontology* (IAO) e *Relation Ontology* (RO); tal que os nomes fora dos parênteses correspondem aos atributos dessas ontologias e os nomes entre parênteses são seus sinônimos no Protégé. As demais atividades dessa quinta etapa da metodologia são: (ii) criação de uma definição formal para cada classe e relação da ontologia: a partir da definição textual de cada classe e relação, elaborada na etapa de conceitualização, constrói-se uma definição formal usando relações ontológicas para relacionar os termos e substituir os verbos incluídos na definição textual; (iii) construção de

outros axiomas formais além das definições formais e propriedades das classes, como forma de assegurar um bom grau de axiomatização da ontologia em desenvolvimento; (iv) definição de instâncias (particulares ou individuais) para cada classe da ontologia, respeitando as restrições definidas para cada classe.

A **sexta etapa** da metodologia proposta corresponde à implementação da ontologia em uma linguagem formal e computável, de forma que o conteúdo ontológico possa ser interpretado pelas máquinas. Optou-se por separar essa etapa (implementação) da etapa de formalização (etapa anterior) pelo fato de que a formalização da ontologia deve ser feita de maneira independente e transparente em relação à sua representação em linguagens formais. Como passos dessa etapa têm-se: (i) a seleção de uma linguagem formal, geralmente, linguagens lógicas, para implementação da ontologia. Em domínios biomédicos as linguagens de lógica descritiva tem sido preferíveis em relação às linguagens de primeira-ordem; (ii) a geração propriamente dita da representação formal da ontologia na linguagem escolhida, geralmente, usando um editor de ontologias que possibilita essa geração automática.

A **sétima etapa** corresponde à avaliação do conteúdo ontológico, que envolve como atividades as seguintes: (i) verificar e tratar erros de inconsistências, incompletudes e ambiguidades dos termos da ontologia, utilizando os classificadores (*reasoners*) disponíveis nos atuais editores de ontologia, tais como FaCT++ e HermiT do Protégé; (ii) especificar questões de competência de validação da ontologia, tanto em linguagem natural quanto em linguagem formal; (iii) utilizar critérios de avaliação da qualidade da ontologia, disponíveis no método formal de validação conhecido como *OntoLingua* (ONTOLINGUA, 2014) e também em outras metodologias de construção, como a *Methontology*; (iv) consultar especialistas no domínio tratado para validação do conteúdo ontológico, apresentando-lhes a ontologia a partir de uma interface de visualização da mesma.

Por fim, tem-se a **oitava etapa** da metodologia que se refere à documentação da ontologia desenvolvida nas etapas anteriores. Nessa documentação devem-se incluir os seguintes elementos essenciais: (i) elementos de especificação da ontologia (propósito geral, escopo de cobertura, público-alvo, etc.); (ii) questões de competência para definição do escopo e também para validação da ontologia; (iii) relação e breve descrição das ontologias reutilizadas no projeto ontológico; (iv) relação dos termos reaproveitados dessas ontologias; (v) dicionários de classes; (vi) dicionário de relações; (vii) hierarquias e outras estruturas de representação da ontologia; (viii) conjunto de critérios, métodos e técnicas usados na avaliação e validação; (ix) relação e descrição dos materiais de referência utilizados na aquisição do conhecimento.

5 CONCLUSÕES E TRABALHOS FUTUROS

O presente artigo apresentou uma revisão e análise das atuais metodologias de construção de ontologias, incluindo nessa análise a sugestão de alguns aspectos ontológicos essenciais que, de alguma forma, devem ser tratados e incluídos em uma metodologia para a construção de ontologias. A realização dessas tarefas possibilitou à presente pesquisa propor um novo conjunto de princípios metodológicos para a construção de ontologias que leva em consideração os aspectos ontológicos abordados e detalha passos do processo de desenvolvimento, em geral, não explicitados na maioria das metodologias atuais.

Como mencionado anteriormente, a elaboração de metodologias próprias para a construção de ontologias, fazendo adaptações as atuais e combinando o que há de melhor em cada uma, tem se tornado uma prática comum na área de engenharia ontológica, como demonstrado por Santos (2014). Pode-se discutir se a elaboração de uma metodologia própria para construção de uma ontologia específica é benéfica ou não para a área de engenharia ontológica, entretanto, essa discussão é bastante ampla e foge ao escopo pretendido pelo presente artigo.

Nesse ponto é importante ressaltar que a metodologia proposta justifica-se por representar uma solução alternativa aos problemas enfrentados às atuais metodologias e também porque servirá como instrumento de apoio aos desenvolvedores de ontologias em domínios biomédicos relacionados ao sangue, inseridos em um projeto nessa área denominado de *Blood Project*. Tal projeto objetiva a representação formal de todo o domínio do sangue, para o qual ainda não se conhece uma ontologia específica sobre o tema, incluindo uma série de ontologias para tal (ALMEIDA *et al.*, 2010).

O uso da metodologia proposta no escopo do *Blood Project*, por diferentes desenvolvedores de ontologias, servirá como uma boa base de testes da mesma, colocando em prática os princípios e fundamentos ontológicos aos quais essa metodologia está baseada. Atualmente, tal metodologia já vem sendo usada no desenvolvimento de uma ontologia sobre hemocomponentes e hemoderivados do sangue humano, denominada HEMONTO (MENDONÇA e ALMEIDA, 2013), que é uma das ontologias inseridas no âmbito do *Blood Project*.

Levando-se em consideração que a elaboração de metodologias próprias é uma das atividades de pesquisa que, de alguma forma, contribuem para o desenvolvimento científico e que a aplicação prática de tal metodologia envolve um caráter social e humano relevantes, devido a representação do conhecimento no domínio do sangue, acredita-se que tal pesquisa

vem a contribuir com um artefato de representação e organização da informação, comuns nas pesquisas em Ciência da Informação.

REFERÊNCIAS

- ALMEIDA, M. B.; TEIXEIRA, L. M. D.; COELHO, K. C.; SOUZA, R. R. Relações semânticas em ontologias: estudo de caso do Blood Project. **Liinc em Revista**, v.6, n.2, setembro, 2010, Rio de Janeiro, p. 384- 410.
- BERNARAS, A.; LARESGOITI, I.; CORERA, J. Building and reusing ontologies for electrical network applications. In: THE EUROPEAN CONFERENCE ON ARTIFICIAL INTELLIGENCE, ECAI, 1996. **Proceedings...** Ecai, 1996. p. 298-302.
- BITTNER, T.; DONNELLY, M. Logical properties of foundational relations in bio-ontologies. **Artificial Intelligence in Medicine**, v. 39, n. 3, p. 197–216, 2007.
- BREITMAN, K. **Web semântica: a Internet do futuro**. Rio de Janeiro: LTC, 2005.
- CARDOSO, J. The semantic web vision: where are we?. **IEEE Intelligent Systems**, p. 22-26, Sept./Oct. 2007.
- DE NICOLA, A.; MISSIKOFF, M.; NAVIGLI, R. A software engineering approach to ontology building. **Information Systems** 34, p. 258–275, 2009.
- DEGEN, W.; HELLER, B.; HERRE, H.; SMITH, B. **GOL: toward an axiomatized upper-level ontology**. In: 2nd INT. CONF. ON FORMAL ONTOLOGY IN INFORMATION SYSTEMS, 2., 2001. **Proceedings...** New York, USA: ACM, 2001. p.34–46. Ogunquit, USA, 2001.
- FERNÁNDEZ, M. *et al.* Building a chemical ontology using methontology and the ontology design environment. **Intelligent Systems**, v. 14, n. 1, p. 37-46, Jan./Feb.1999.
- FERNÁNDEZ, M.; CORCHO, O. Methodologies and methods for building ontologies. In: FERNÁNDEZ, M.; CORCHO, O. **Ontological engineering**. London: Springer, 2004. p. 107-153.
- GÓMEZ-PÉREZ, A.; FERNÁNDEZ, M.; VICENTE, A. J. **Towards a method to conceptualize domain ontologies**. In: ECAI WORKSHOP ON ONTOLOGICAL ENGINEERING, 1996, Budapest. Disponível em: <<http://citeseer.ist.psu.edu/483876.html>>. Acesso em: 13 de março de 2014.
- GRENON, P.; SMITH, B. SNAP and SPAN: towards dynamic spatial. **Spatial Cognition & Computation**, v.4, n.1, p. 69-104, 2004. Disponível em: <http://ontology.buffalo.edu/smith/articles/SNAP_SPAN.pdf>. Acesso: 12 de Abril de 2012.
- GRUNINGER, M.; FOX, M. S. **“Methodology for the Design and Evaluation of Ontologies”**. In: WORKSHOP ON BASIC ONTOLOGICAL ISSUES IN KNOWLEDGE SHARING (IJCAI). [S.l.; s.n.], 1995.
- GUARINO, N. Formal Ontology in Information Systems. **Proceedings of FOIS'98**, Trento, Italy, 6-8 June 1998. Amsterdam, IOS Press, pp. 3-15, 1998.
- HERRE, H.; HELLER, B.; BUREK, P.; HOEHNDORF, R.; LOEBE, F.; MICHALEK, H. **General Formal Ontology (GFO): A Foundational Ontology Integrating Objects and**

Processes. Part I: Basic Principles. Leipzig: Institute of Medical Informatics, Statistics and Epidemiology (IMISE), p. 16 - 18, 2006.

JONES, D.; BENCH-CAPON, T.; VISSER, P. **Methodologies for ontology development.** 1998. Disponível em: <<http://cweb.inria.fr/Resources/ONTOLOGIES/methodo-for-ontodev.pdf>>. Acesso em: 13 de março de 2014.

LIM, S. C. J.; LIU, Y.; e LEE, W. B. A methodology for building a semantically annotated multi-faceted ontology for product family modelling. **Advanced Engineering Informatics** 25 (2), pp. 147-161, 2011.

MASOLO, C.; BORGIO, S.; GANGEMI, A.; GUARINO, N.; OLTRAMARI, A. **Ontology Library: WonderWeb Deliverable D18.** Trento, Italy, 2003. Disponível em: <<http://www.loa-cnr.it/Papers/D18.pdf>>. Acesso em: 05 de agosto de 2012.

MENDONÇA, F.M.; ALMEIDA, M.B. (2013). Hemocomponents and Hemoderivatives Ontology (HEMONTO): an ontology about blood components. Disponível em:

<http://mba.eci.ufmg.br/downloads/ontobras-2013_paper1.pdf>. Acesso em: 21 de abril de 2014.

NOY, N. F.; McGUINNESS, D. L. **Ontology Development 101: A Guide to Creating Your First Ontology.** Disponível em:

<<http://www.ksl.stanford.edu/people/dlm/papers/ontology101/ontology101-noymcguinness>> Acesso em: 13 de março de 2014.

OBO - Open Biomedical Ontologies. Disponível em: <<http://www.obofoundry.org>>. Acesso em: 4 de Maio de 2014.

REED, S.L.; LENAT, D.B. **Mapping Ontologies into Cyc,** 2002. Disponível em: <http://www.cyc.com/doc/white_papers/mapping-ontologies-into-cyc_v31.pdf>. Acesso em: 25 de abril de 2002.

SANTOS, M. T. **Estudo do processo de apropriação da ontologia pela Ciência da Informação no Brasil.** Dissertação (Mestrado), Universidade Federal de Pernambuco, Recife, PE, Brasil, 2014.

SCHULZ, S.; STENZHORN, H.; BOEKER, M.; e SMITH, B. Strengths and limitations of formal ontologies in the biomedical domain. **RECIIS – Electronic Journal Communication Information & Innovation in Health.** Rio de Janeiro, v.3, n.1, 31-45, Mar., 2009.

SCHIESSL, M.; BRÄSCHER, M. Do texto às ontologias: uma perspectiva para a ciência da informação. **Ciência da Informação,** Brasília, v. 40, n. 2, maio/ago. 2011.

SILVA, D. L.; SOUZA, R. R.; e ALMEIDA, M. B. Ontologias e vocabulários controlados: comparação de metodologias para construção. **Ciência da Informação,** Brasília, v. 37, n. 3, p. 60-75, set./dez. 2008.

SMITH, B. (2013). Introduction to the Logic of Definitions. Disponível em: <<http://www.slideshare.net/BarrySmith3/logic-ofdefinitions>>. Acesso em: 12 de maio de 2014.

SMITH, B. **Ontology and Information Systems**, 2004. Disponível em:
<<http://www.ontology.buffalo.edu/ontology>> Acesso em: 16 de março de 2012.

SMITH, B.; CEUSTERS, W.; KLAGGES, B.; KÖHLER, J.; KUMAR, A.; LOMAX, J.; MUNGALL, C.; NEUHAUS, F.; RECTOR, A. L.; ROSSE, C. Relations in biomedical ontologies. **Genome Biology** 6, R46, 2005.

SMITH, B. et al. The OBO foundry: Coordinated evolution of ontologies to support biomedical data integration. **Nature Biotechnology**, v. 25, n. 11, p. 1251–1255, 2007.

SOERGEL, D. **Functions of a Thesaurus / Classification / Ontological Knowledge Base**. College of Library and Information Services, University of Maryland, 1997.

SUÁREZ-FIGUEROA, M. C. NeOn Methodology for Building Ontology Networks: Specification, Scheduling and Reuse. Tese (Doutorado), 2010.

SURE, Y.; STAAB, S.; e STUBER, R. On-To-Knowledge Methodology (OTKM). 2003. Disponível em:
http://www.sfu.ca/~mhatala/iat881/papers/2003_ontohandbook_otkmethodology.pdf. Acesso em: 17 de março de 2014.

SWARTOUT, B. et al. **Toward distributed use of large-scale ontologies**. 1996. Disponível em: <http://ksi.cpsc.ucalgary.ca/KAW/KAW96/swartout/Banff_96_final_2.html>. Acesso em: 13 de março de 2014.

USCHOLD, M.; KING, M. **Towards a methodology for building ontologies**. 1995. Disponível em: <<http://citeseer.ist.psu.edu/uschold95toward.html>> Acesso em: 11 mar. 2014.

USCHOLD, M.; GRUNINGER, M. Ontologies: Principles, methods and applications. **Knowledge Engineering Review**, v. 11, n. 2, p. 93–136, 1996.

VICKERY, B. C. Ontologies. **Journal of Information Science**, v. 23, n. 4, p. 277-286, 1997.

**A ANÁLISE TIPOLOGICA COMO SUBSÍDIO PARA CONSTRUÇÃO DE
VOCABULÁRIO CONTROLADO: O CASO DO JUÍZO DE ÓRFÃOS DO RECIFE
(1824-1889)**

**THE TYPOLOGICAL ANALYSIS AS SUBSIDY FOR CONSTRUCTION OF
CONTROLLED VOCABULARY: THE CASE OF RECIFE ORPHAN'S COURT (1824-
1889)**

Mônica Maria Pádua Souto da Cunha
Fábio Mascarenhas Silva

Resumo: O objetivo desta pesquisa foi desenvolver um percurso metodológico, unindo-se as técnicas da Tipologia Documental e da Análise Documentária, orientado à compreensão contextual da produção do acervo do Juízo de Órfãos dos oitocentos, a partir da análise comparativa entre processos judiciais similares, mas produzidos em épocas distintas. As balizas temporais para esta pesquisa foram determinadas pela data da criação do Juízo de Órfãos pela legislação brasileira, depois da grande reforma do Judiciário a partir da Constituição de 1824, até a Proclamação da República (1889). O referencial teórico apresentou conceitos, histórico, características, modelos e experiências sobre as temáticas, diplomática, análise tipológica em arquivos e a interseções entre elas, e análise documentária e vocabulário controlado. Na primeira fase do desenvolvimento da metodologia, reconheceu-se o conteúdo informativo do conjunto documental objeto do estudo, a partir de métodos da Tipologia Documental e da Arquivologia. Na segunda fase, identificaram-se as características dos documentos propriamente ditos, aplicando-se os conhecimentos da Análise Documentária, em especial do vocabulário controlado em arquivos. Percebeu-se que, para melhor recuperar as informações contidas na documentação arquivística de caráter permanente é preciso congrega metodologias distintas.

Palavras-chave: Diplomática, Análise tipológica, Análise documentária, Juízo de órfãos, Memorial da Justiça de Pernambuco.

Abstract: This research aims develop a methodological approach, combining the techniques of Typological Analysis and Documentary Analysis, oriented contextual understanding of the production of the acquits of Judgment Orphans of eight, from the comparative analysis Similar lawsuits, but produced at different times. The temporal boundaries for this study were determined by the date of creation of the Brazilian Orphans Court, after the great reform of the judiciary from the 1824 Constitution until the Republic Proclamation (1889). The presented theoretical concepts, historical, features, models and experiences on issues, diplomatic, typological analysis on archives and intersections between them, and documentary analysis and indexing language. In the first stage of the development methodology, it was recognized the information content of the set of documents under study from Typological Analysis and methods Archival. In the second phase, we identified the characteristics of the documents themselves, applying knowledge of Documentary Analysis, especially in the controlled vocabulary files. We realized that to better recover the information contained in the permanent archive documentation is necessary to bring together different methodologies.

Key-words: Typological Analysis, Diplomatic, Indexing languages, Brazilian Orphans Court, Pernambuco Justice Memorial

1 INTRODUÇÃO

O Tribunal de Justiça de Pernambuco (TJPE) criou, desde 1998, o Memorial da Justiça, que atualmente guarda mais de vinte conjuntos documentais de valor permanente das comarcas de Pernambuco, com períodos cronológicos que variam desde meados do século XVIII até o início do século XX. A maior parte desse acervo se encontra identificada por um Guia de Fundos, instrumento voltado à recuperação de documentos e contém as características descritivas dos acervos históricos, tais como o título do fundo, o período cronológico de cada acervo, a quantidade de documentos de cada conjunto documental. Aos usuários também está aberto à consulta o inventário sumário dos acervos, que oferece outras informações em formato de listagem, como o nome das ações ou recursos de cada caixa que guarda a documentação.

A lista de termos, contendo o nome da ação ou recurso, disponível em inventário sumário no Memorial, foi elaborada a partir da análise individual de cada documento, da capa do processo ou das informações coletadas em uma de suas peças. Essa estratégia de identificação documental foi definida em caráter emergencial para início das atividades do Memorial, no intuito de disponibilizar ao público o acervo, já que um processamento técnico mais complexo, envolvendo a indexação documental demandaria um maior tempo, sobretudo na identificação da terminologia jurídico do período de criação do processo, bem como as características internas e externas de cada dossiê.

Contudo, a indexação de cada processo judicial, sem prévio conhecimento sobre a matéria do acervo, gerou dúvidas ao corpo técnico do Memorial. Essas dúvidas decorreram, na maior parte dos casos, de processos judiciais identificados por termos em desuso, que remontam às práticas jurídicas dos séculos XVIII e XIX, pautadas em leis já revogadas.

Segundo levantamento feito em 2012 (CUNHA, 2012), os usuários do Memorial da Justiça, de 2007 a 2011, pesquisaram em processos judiciais relacionados ao patrimônio das pessoas, como é o caso do Juízo de Órfãos do século XIX, e que 29% das pesquisas no acervo do Memorial centravam-se nos órgãos cujas atividades assemelharam-se àquelas de competência do Juízo de Órfãos. Essa foi a motivação para se estudar esse conjunto documental e desenvolver este estudo.

Conforme afirma Cardozo (2010), criou-se o Juízo de Órfãos no Brasil para regulamentar as relações entre as famílias constituídas com menores de idade. Foi o tribunal ou foro no qual se decidiam questões relativas aos menores de idade ou equiparados. Inicialmente cuidou das questões relativas às partilhas, inventários, heranças e guarda dos

menores da elite. Ao longo do tempo suas atribuições foram ampliadas, segundo a demanda do Estado e da sociedade.

Desta forma, o objetivo desta pesquisa é desenvolver um percurso metodológico voltado à organização da informação, unindo as técnicas da Análise tipológica (AT) e da Análise Documentária (AD), orientado à contextualização da formação do acervo do Juízo de Órfãos a partir da análise comparativa entre processos judiciais similares que foram produzidos em épocas diferentes. As balizas temporais para esta pesquisa foram determinadas pela data da criação do Juízo de Órfãos pela legislação brasileira, depois da grande reforma do Judiciário na Constituição Brasileira de 1824, que se deu principalmente com a publicação da Lei de 29 de novembro de 1832 (primeiro Código de Processo Criminal Brasileiro, com disposição provisória acerca da administração da Justiça Civil).

2 REFERENCIAL TEÓRICO

Nesta seção serão discutidas, sucintamente, as temáticas conceituais que subsidiaram o arcabouço teórico-metodológico desta pesquisa: a diplomática, a análise tipológica, e a análise documentária.

2.1 A Diplomática

Como informa Duranti (1995, p.4-7), a Diplomática foi criada para suprir a necessidade de analisar, criticamente, documentos que poderiam ser falsificados. Documento, para essa ciência, é aquele escrito que se produz em um suporte, por meio de um instrumento. Esse seria, mais especificamente, o documento arquivístico, aquele criado por uma pessoa física ou jurídica no curso de uma atividade administrativa. Herrera (1988, p.37) conceitua Diplomática como “a ciência que estuda o documento, sua estrutura e cláusulas, para estabelecer as diferentes tipologias e sua gênese nas instituições, com o fim de analisar sua autenticidade”.

Importante salientar a visão de Bellotto (2008, p.1), que destaca a insuficiência do conceito simplista e redutor que se tinha da Diplomática, conhecida “como instrumental de fundo paleográfico, historiográfico e jurídico, apto apenas a fornecer a descrição e a explicação da estrutura formal dos atos escritos, sua autenticidade e fidedignidade”. Esse tipo de análise documental foi utilizado largamente nos séculos XVII e XVIII pelo Direito e pela História, para verificar se um documento era autêntico e original, mas atualmente ele se alargou, diante da utilização da Diplomática pela Arquivologia.

O uso da Diplomática pela Arquivologia faz daquela antiga ciência uma ciência renovada. Deu-se início a essa etapa no final do século XX, período em que os arquivistas

começaram a desenvolver pesquisas, usando os princípios e métodos da Diplomática para compreender melhor os processos de criação dos documentos da burocracia moderna. (RONDINELLI, 2005, p.45).

Saliente-se que a fusão dos princípios e métodos arquivísticos e diplomáticos ocorreram na Itália, a partir dos estudos de Carucci sobre a matéria, com a publicação do II documento contemporâneo: *Diplomática e criteri di edizione*. Nesse trabalho a autora define Diplomática como a disciplina que estuda o documento e o arquivo sob a ótica dos aspectos formais, com vistas a definir a natureza jurídica dos atos, tanto na sua formação, quanto em seus efeitos (CARUCCI, 1987).

A metodologia desenvolvida pelos Arquivistas Municipais de Madrid, nos anos 1980, bem como nos manuais de arquivística na Espanha, como é o caso de Cruz Mundet, em 2001, serviram de referência para o desenvolvimento de trabalhos sobre análise diplomática e Tipologia Documental no Brasil (RODRIGUES, 2008).

2.2 Os princípios da Arquivística e o processo de análise tipológica em arquivos

A informação arquivística se diferencia das informações de outra natureza por ser produzida no exercício das funções administrativas, e de acordo com os objetivos de uma determinada entidade. Por essa razão é que se reconhece o caráter orgânico dessas informações, que devem receber tratamento único e a identificação dos documentos não pode ser realizada isoladamente, com relação ao mesmo fundo documental³² (BELLOTTO, 2002).

Em razão da origem espontânea e natural que se estabelece devido a essa forma diferenciada de produção documental, gerada a partir das atividades e funções exercidas pelas instituições produtoras, foram criados princípios que manterão o respeito a essa origem e ao processo natural de sua existência (HERRERA, 1988).

Bellotto (2002b, p. 23) esclarece que a Análise Tipológica deverá sempre trazer consigo o respeito aos princípios da Arquivística, pois eles estabelecem essas peculiaridades dos documentos de arquivo. São eles: o da *unicidade*, o documento de arquivo é único, pois no contexto de produção, no momento da sua criação, não há outro que lhe seja idêntico em propósito pontual, nem em seus efeitos; o da *organicidade*, os arquivos são o espelho da estrutura, funções e atividades da entidade que o produziu e acumulou; o da *indivisibilidade*, consequência do princípio da proveniência, determina que um fundo não possa ser separado

³² Conjunto de documentos produzidos e acumulados por uma determinada pessoa, física ou jurídica, no exercício de suas atividades ou funções.

nem misturado a outros; o da *proveniência*, que para Schellenberg (2006), é aquele princípio de agrupar os documentos oficiais de acordo com a origem nas instituições que os criaram.

Seria possível, então, conceituar documento de arquivo de forma diferente, sem se utilizar dos princípios, ressaltando os dois polos: de um lado, a matéria/suporte; de outro, o conteúdo/informação. Mas esse tipo de identificação não é suficiente. Para o conceito tornar-se completo requer-se retratar a estrutura do documento, composta por elementos extrínsecos e intrínsecos. Os extrínsecos (físicos, de estrutura ou formais) têm relação com o gênero, a forma e os sinais usados para registrar a mensagem; os intrínsecos (substantivos ou de substância) relacionam-se ao conteúdo, ao assunto, à proveniência, à função (BELLOTTO, 2002).

Conforme aborda Rodrigues (2008), em se tratando de um método analítico de investigação, a metodologia diplomática, que se fixa nos procedimentos administrativos de criação dos documentos, e considera também as estruturas organizacionais e suas competências. O estudo do documento na análise diplomática reflete as relações expressas entre a forma e o contexto em que foi produzido.

Esse novo tipo de abordagem da Diplomática chamada de Tipologia Documental ou Análise Tipológica traz para a Arquivística a possibilidade de entender a ação da qual participa o documento, identificando-o em qualquer tipo de sistema. Esse método passou a representar os princípios da Arquivologia, “incorporando novos elementos para identificar o documento no interior do conjunto ao qual pertence” (RODRIGUES, 2008, p. 150-151).

2.3 Análise tipológica, análise diplomática: relações e aplicações

A análise diplomática auxilia a identificação dos tipos documentais arquivísticos. Isto é percebido nesta pesquisa, pois a Análise Tipológica subsidiou a compreensão das características de cada documento produzido pelo Juízo de Órfãos da Comarca do Recife, no século XIX, permitindo o reconhecimento de diversos tipos de processo existentes durante o período cronológico estudado, e a estabelecer as suas diferenças, por meio da representação (CUNHA, 2013). Essa forma de analisar o documento ajuda a delimitar as séries no momento da classificação nos arquivos, confirmando o nosso entendimento, mantendo sua unidade, organicidade e continuidade.

Constata-se a importância da Análise Tipológica para a identificação das séries documentais. Tais séries são efetivamente adotadas na classificação dos acervos, bem como no processo de avaliação, seleção e destinação dos documentos e de acesso dos usuários. Como afirma Carucci (1987), para identificar as séries é preciso analisar as características

formais (itens como número do documento, assinatura, data, nome do autor, nome do órgão que produziu o documento) e substanciais dos documentos, a fim de se agrupar na mesma série os documentos similares no que tange ao procedimento, independentemente de serem denominados diferentemente.

Herrera (1988) lembra que os tipos documentais são gerados por atividades praticadas dentro das instituições, no desenrolar de determinadas funções. Para ela, a delimitação dos tipos documentais dependerá da análise dos caracteres externos e internos dos documentos e da informação contida neles. Para a fixação dos termos utilizados para cada tipo de documento, sejam jurídicos ou administrativos, que modificam com as mudanças sociais (e legais) de cada época, exige-se um estudo pormenorizado no qual a metodologia da análise diplomática é útil.

Fixando-se nos procedimentos administrativos de criação dos documentos nas estruturas organizacionais, e considerando suas competências, o método de análise diplomática fundamenta a identificação das funções, essencial para a compreensão dos documentos dentro do seu ciclo de vida (RODRIGUES, 2008). A partir daí o arquivista pode organizar os documentos em séries, com a identificação das atividades e classificá-los pela semelhança entre funções que os criaram. Assim foi realizado o trabalho com os processos judiciais do Juízo de Órfãos da Comarca do Recife. Primeiro, analisou-se a estrutura organizacional a partir da qual foi produzido e acumulado o acervo, suas competências determinadas pela legislação do período, momento no qual foram analisadas suas funções em cada um dos momentos históricos estudados (CUNHA, 2013).

2.4 Modelos de Análise Tipológica

Rondinelli (2005) afirma que analisar um documento arquivístico, com base na metodologia diplomática, significa pressupor uma relação entre a palavra e o mundo. Para tanto, faz-se necessário decompor o documento em seus elementos constitutivos, suporte, conteúdo, forma ou estrutura, ação, agentes geradores ou pessoas, relação orgânica e contexto, buscando compreender o mundo por meio dele.

Sabe-se que a metodologia de tratamento documental para a Diplomática relaciona-se com o verídico, com a estrutura e a finalidade do ato jurídico. Explica Bellotto (2002) que para a Tipologia Documental o que importa é a relação dos documentos com as atividades institucionais/pessoais. Ainda segundo a mesma autora, a Diplomática e a Arquivística decodificam o documento de formas diferentes, mas no final chegam ao mesmo resultado: o

conhecimento da forma (suporte, meio, contextualização) e a sua relação com o seu conteúdo (aquilo que se pretende ao emitir-se o documento).

Sobre modelos de Análise Tipológica utilizados na Arquivologia, citam-se alguns importantes registros encontrados na literatura. O Grupo de Trabalho dos Arquivistas Municipais de Madri elaborou o Manual de Tipologia Documental estabelecendo parâmetros fornecidos por essa análise para possibilitar a avaliação e seleção documental (GAGNON-ARGUIN, 1998). Esse manual é considerado como precursor dos trabalhos de Tipologia Documental na área da Arquivologia, destaca-se por suas contribuições no processo de construção do vocabulário controlado.

O modelo criado pelos Arquivistas de Madri volta-se aos arquivos correntes e parte dos princípios da Arquivística, considerando os seguintes itens: tipo (espécie documental + atividade concernente); código da série correspondente ao tipo no plano de classificação; entidade produtora acumuladora com subdivisões que correspondem às funções; atividades que geram o tipo documental; destinatário; legislação que cria a entidade; tramitação (sequência de procedimentos); documentos básicos que compõem o processo; ordenação dentro da série; conteúdo; vigência; prazos de permanência em cada arquivo setorial, eliminação ou preservação (BELLOTTO, 2002).

Outro modelo de Análise Tipológica é aquele elaborado pela autora canadense Gagnon-Arguin. Similar à metodologia exposta no Manual supracitado, difere por identificar as “espécies documentais que podem ser encontradas nas entidades privadas ou nos órgãos estatais com suas características sem, entretanto, fazer com que sobre ela incida a exemplificação direcionada” (BELLOTTO, 2002, p.99-100).

No Brasil, Bellotto, inspirada na experiência desenvolvida pelo Grupo de Arquivistas Municipais de Madri, inova, apresentando os fundamentos da Tipologia Documental, em 1980 e 1990 (RODRIGUES, 2008). Guimarães e Nascimento (2007) desenvolvem, na década de 1990, trabalho metodológico principalmente com a aplicação da Diplomática na indexação de jurisprudência, acreditando que essa ciência oferece um referencial teórico capaz de auxiliar na identificação das espécies/tipologias dos documentos, por meio da partição com a análise de sua estrutura.

Duranti, no Canadá, em 1993, integrante de um projeto de pesquisas sobre documentos digitais, constrói um modelo de análise elencando todos os elementos que caracterizam um documento arquivístico. A autora recomenda que se aplique a metodologia diplomática da seguinte forma: definindo o documento, identificando seus componentes e elementos detalhadamente e interpretando suas funções de forma aprofundada, tudo dentro do

contexto do tipo de entidade produtora e das funções por ela desempenhada, indo além da Diplomática clássica. Ela contribui introduzindo um novo caminho para a análise dos tipos documentais, uma forma de se “comparar documentos produzidos em períodos, locais e por órgãos diferentes, com finalidades diversas, por meio da identificação dos elementos extrínsecos e intrínsecos que compõe a estrutura do documento” (RODRIGUES, 2008, p.145-147).

Com a publicação da obra intitulada *Diplomatics: new uses for na old science*, Duranti se destaca nos estudos sobre a aplicação do método diplomático pela Arquivologia, confirmando em seus estudos que a Análise Tipológica por meio da comparação, pode ser considerada uma nova forma de compreender os documentos contemporâneos.

2.5 Indexação e vocabulário controlado como recursos da análise documentária

Para Lancaster (2004) indexar é identificar o conteúdo do documento a partir da atribuição de termos, que servem como pontos de acesso para localizar e recuperar uma informação. Kobashi (2008, p. 2) afirma que indexar é caracterizar conteúdos de documentos por meio dos descritores de um vocabulário controlado. Ler e interpretar textos é uma das operações fundamentais da indexação. A leitura consiste em captar os significados essenciais do conteúdo informacional de documentos a fim de representá-los adequadamente para as funções de identificação e localização de informação pelo usuário. A leitura em documentação tem por objetivo representar conteúdos informacionais para o controle e recuperação de documentos.

Evidencia-se a indexação como caminho para a criação de vocabulário controlado, e este como um dos alicerces de conhecimento para o desenvolvimento desta pesquisa, apresentando-se como uma base metodológica eficiente para a análise do acervo do Juízo de Órfãos no século XIX.

Frisa-se que a documentação processual cível, da qual faz parte o acervo analisado neste trabalho, apresenta especificidades da área da qual é integrante: o Direito Processual Civil. Nessa área existem procedimentos judiciais que podem ser representados por classes, e estas, por sua vez, desdobradas e interligadas hierarquicamente. Essas classes processuais, que são as tipologias documentais (tipos de ação e recurso, segundo a tabela processual unificada de classes do CNJ³³), relacionam-se a conteúdos que são representados por assuntos (de

³³ Disponível em:<http://www.cnj.jus.br/sgt/consulta_publica_classes.php>. Acesso em: 10 jun. 2014.

acordo com a tabela processual unificada de assuntos do CNJ³⁴). Os assuntos, nesse caso, são a consequência de cada uma das funções para as quais os documentos foram criados, a partir da legislação.

Kobashi (2008) conceitua vocabulário controlado como uma linguagem artificial, composta por termos organizados e relacionados entre si, utilizada para indexar documentos. Tem como finalidade a padronização, com vistas a facilitar a entrada e a saída de dados em um sistema de informações, permitindo uma comunicação mais precisa e eficaz entre os usuários e o sistema de informações.

Para o acervo do Juízo de Órfãos dos oitocentos, documentação de caráter arquivístico, pode ser utilizada a solução sugerida por Smit (2003) para elaborar vocabulário controlado. A autora inicia seu manual conceituando a série documental como “sequência de unidades de um mesmo tipo documental”, explicando que os documentos de uma mesma tipologia se agregam por apresentar uma característica comum capaz de diferenciá-los de outros tipos documentais, com a consequente exclusão dos outros diferentes. Conclui que essa organização pressupõe a identificação de semelhanças e de diferenças, que se caracterizam como uma forma de ponto de acesso, e este serve de mediador para que se dê o acesso nos arquivos (SMIT, 2003, p. 11-12).

3 DESENVOLVIMENTO DA METODOLOGIA PROPOSTA

3.1 Os passos do trabalho

Com vistas a definir os procedimentos de análise dos documentos do *corpus* deste estudo, exigiu-se um prévio conhecimento do documento em suas características e objetivos (GUIMARÃES, 1993). Por essa razão, optou-se por, inicialmente, desenvolver um estudo sobre o contexto no qual o acervo foi produzido e acerca do fim para o qual foram criados os documentos.

Nessa primeira fase foi imprescindível conhecer o conteúdo informativo do conjunto documental objeto do estudo, nesse caso, o Juízo de Órfãos da Comarca do Recife, suas funções e organização interna e as relações entre seus documentos, bem como entender quais as atividades que as geraram, pois, segundo Duranti (1995), “se o arquivo é um todo constituído por partes, é impossível entender e controlar o todo sem compreender e controlar suas partes, ainda que as mais elementares”. Para isso, foi preciso analisar a legislação

³⁴ Disponível em: <http://www.cnj.jus.br/sgt/consulta_publica_assuntos.php>. Acesso em 10 de junho de 2014.

aplicada na época imperial ao Juízo de Órfãos, em especial a Constituição de 1824, a Lei de 29 de novembro de 1832 (primeiro Código Criminal Brasileiro) a Lei do Ventre Livre (Lei nº 2.040 de 28 de setembro de 1871), bem como outras fontes sobre o assunto e acerca do século XIX no Brasil e em Pernambuco, produzidas nas áreas do Direito, citem-se Silveira (1876), Carvalho (1879), Cerne (2011) e Beviláqua (1946, 1978); na História do Direito, como Malheiro (1866) e Ferreira (1956); e na História sobre o Império, na qual podem se citadas as obras de Lyra (1941) e Neves (1999).

Somando-se à leitura das fontes legais e literatura relacionadas ao período cronológico da pesquisa, foi importante, para compreender o conteúdo informativo do acervo em estudo, o manuseio dos documentos. Essa ação foi complementar, mas imprescindível para se entender as relações existentes entre as atividades do órgão e seus documentos, construindo-se, assim, o alicerce para a sua identificação.

Importante lembrar que o Direito é uma área especializada e, devido a isso, os documentos produzidos a partir das atividades específicas dessa área têm características próprias, com funções bem definidas e usuários determinados. Como bem informa Guimarães,

a forma documentária do Direito varia em função do objetivo do próprio documento [...] variáveis como o produtor da informação, o grau de coercibilidade e o fim a que se destina serão de fundamental importância seja na sua produção, seja no seu tratamento temático (GUIMARÃES, 1993, p. 53).

Pela razão de que o Direito é uma área técnica e sua forma documentária é bem característica, haja vista que um processo judicial somente existe como consequência da prestação jurisdicional de órgão judicial, houve a necessidade de aprofundar esse estudo com base na legislação da época para compreender o universo analisado neste trabalho.

Após a primeira etapa, identificaram-se as características dos documentos propriamente ditos. Nesse caso, identificação é “o ato de caracterizar os caracteres próprios e exclusivos que conferem a identidade do documento de arquivo”, tem como consequência a individualização de um documento em relação ao seu conjunto, pois determina os elementos que o distinguem do todo (RODRIGUES, 2008, p. 65).

Para normalizar os termos identificadores das ações e recurso judiciais produzidos pelo Juízo de Órfãos da Comarca no Recife, acataram-se as recomendações e técnicas dos princípios da proveniência, por definir a organicidade dos fundos e dos arquivos, determinando a sua classificação (HERRERA, 2003) e da ordem original, por indicar a ordem sequencial dos documentos no conjunto que integram, refletindo o procedimento do qual se

originou o documento (HERRERA, 2003) - pois está em foco um acervo arquivístico -, em conjunto com a base metodológica da Tipologia Documentária e da Análise Documental. A relação entre essas metodologias favorece a construção de padrões que possam servir como base para a elaboração de instrumentos de busca mais precisos, com vistas à ampliação do acesso relacionado a acervos arquivísticos, caso da documentação estudada.

Na fase de identificação dos documentos, foram aplicados os princípios da proveniência e o da ordem original, justificando-se para isso o fato de que o primeiro define a organicidade dos fundos e dos arquivos, determinando a sua classificação, enquanto o segundo indica a ordem sequencial dos documentos no conjunto que integram e, por consequência, reflete o procedimento do qual se originou cada documento (HERRERA, 2003). Tais ações foram necessárias para entender o conjunto documental em questão, sua formação e as relações entre os seus documentos, pois seria difícil realizá-lo sem compreender o funcionamento daquela instituição e as relações das suas funções com a produção documental existente.

A metodologia apresentada por Rodrigues (2008) em sua tese foi um dos parâmetros para o desenvolvimento da identificação dos documentos nesta pesquisa. Em primeiro lugar, conforme já descrito, é importante buscar informações sobre o órgão produtor em fontes específicas, para se entender o contexto no qual ele foi criado, e as tipologias documentais existentes. Isso se tornou exequível por meio da análise das funções exercidas pelo Juízo Orfanológico, nos oitocentos, determinadas pela legislação, fonte preciosa para se coletar essas informações. Os dados extraídos foram cadastrados em fichários que serviram como instrumento para comparar com as informações contidas na documentação, no momento da identificação e da análise dos processos judiciais.

Como a equipe técnica do Memorial encontrou dificuldade para identificar as peças desse tipo de acervo permanente, de forma a facilitar o acesso ao público, outros arquivos históricos podem também ter enfrentado ou estar enfrentando esse entrave, já que guardam documentos de séculos passados, que apresentam características e termos diferentes dos atuais. Para proceder à identificação desse tipo de acervo, realizou-se uma pesquisa sobre a matéria a partir da qual foram gerados os documentos – o Direito e as leis válidas no período cronológico selecionado, a fim de melhor organizar a informação.

Para construir os termos que deram nome aos processos judiciais foi necessário analisar esse tipo documental com profundidade. Para isso, foram usados os conhecimentos da Diplomática contemporânea, que serviu como um meio valioso para esta identificação tipológica, já que a documentação escolhida para estudo é de arquivo, podendo também

auxiliar na tarefa de arranjo, na medida em que dá subsídios para organização das séries tipológicas, proporcionando outro olhar sobre o documento e viabilizando o desenvolvimento de instrumentos de busca eficientes.

Os procedimentos até aqui apresentados buscaram atingir o objetivo desta pesquisa, que se iniciou com o aporte teórico da Arquivologia para identificar os documentos, pois essa ação garantiu a coleta de informações de forma sistemática, perpassando pela aplicação da metodologia da Tipologia Documental, que assegura o reconhecimento do documento em todas as suas características, viabilizando, por fim, a criação de novos documentos mais resumidos pelo uso dos conhecimentos da Análise Documental, com vistas à construção de um vocabulário controlado, para possibilitar a recuperação da informação, conforme será relatado na próxima seção. Os passos usados para atingir o objetivo deste trabalho estão discriminados no QUADRO 1, com as etapas do percurso metodológico aplicado, com descrição pormenorizada de cada uma das atividades executadas.

QUADRO 1 – Quadro conclusivo: procedimentos metodológicos

ETAPAS	DESCRIÇÃO	ATIVIDADES	ESPECIFICAÇÃO
1	Identificação da legislação (séc. XIX – sobre o Juízo de Órfãos	<ul style="list-style-type: none"> • Identificar a legislação aplicada na época imperial ao Juízo de Órfãos. 	Legislação: em especial a Constituição de 1824, a Lei de 29 de novembro de 1832 (primeiro Código Criminal Brasileiro) a Lei do Ventre Livre (Lei nº 2.040 de 28 de setembro de 1871).
2	Identificação das relações entre as funções/atividades do Juízo de Órfãos com seus documentos produzidos, no contexto histórico e jurídico do período cronológico estudado.	<ul style="list-style-type: none"> • Ler títulos produzidos por juristas do século XIX, bem como literatura sobre o Direito e a História do período imperial brasileiro; • Coletar informações sobre o Juízo de Órfãos. 	<ol style="list-style-type: none"> 1. Títulos produzidos no século XIX e outras fontes: Direito: Silveira (1876), P. P. (1875), Cerne (2009) e Beviláqua (1908, 1946, 1978); História do Direito: Malheiro (1866) e Ferreira (1956); História sobre o Império brasileiro: Lyra (1941) e Neves (1999). 2. Coleta das seguintes informações: data de criação e extinção; data de todos os textos normativos que determinaram uma mudança na estrutura do organismo; órgãos que tenham precedido o desenvolvimento das competências análogas; órgãos que herdaram competências semelhantes; tipo de competência relacionada a cada período cronológico e legislação.
3	Realização de levantamento sobre as características dos documentos do acervo pesquisado, as relações dos documentos com as funções/atividades do Juízo de Órfãos nos oitocentos.	Identificar dos seus elementos intrínsecos, componentes da estrutura do documento, por meio da leitura do acervo selecionado e da comparação entre aqueles produzidos em períodos diferentes.	<ol style="list-style-type: none"> 1. Elementos intrínsecos: protocolo e subseções; texto e subseções; escatocolo e subseções 2. Pessoas: autor da ação; autor do documento; destinatário da ação; destinatário do documento; escritor; contra-assinatura(s) 3. Natureza de ação: processo judicial cível ou criminal 4. Nome da ação: venda, autorização, solicitação (ex.) 5. Relação entre documento e procedimento: especificação da fase do procedimento geral ao qual documento e os documentos se relacionam e, se é resultado de uma “norma de procedimento”; a fase do procedimento específico 6. Tipo de documento: nome; natureza; função; status 7. Descrição diplomática: contexto; ação; documento 8. Comentários conclusivos

4	Identificação dos nomes que a legislação determinava para os documentos.	Comparar o estudo realizado no item 1 e relacioná-los com aqueles encontrados nos próprios documentos, mediante a leitura dos mesmos e os resultados do item 3.	Elaboração de listagem com as informações comuns e diferentes identificadas da legislação e em cada documento analisado, para proceder ao tratamento da informação.
5	Tratamento sistemático da informação, a partir da identificação analítica do conteúdo informativo do documento.	Realizar o levantamento de conceitos que deverão ser traduzidos em expressões representativas de assunto.	Observação da estrutura do documento, conforme preceitua a Diplomática Contemporânea e sua relação com os termos identificados.
6	Realização da seleção dos conceitos identificados.	Diferenciar os conceitos principais dos conceitos secundários dos documentos.	Transformação do documento original em microdocumento (condensação).
7	Realização da representação documental.	Traduzir o conteúdo temático do documento em linguagem de indexação (vocabulário controlado).	Seleção dos termos a serem organizados em vocabulário controlado, com base nos procedimentos indicados por SMIT (2003).
8	Produção de índices.	Definir os termos.	Utilização dos procedimentos micro e macro para definir os termos (SMIT, 2003).

Fonte: elaborado pelos autores

3.2 Análises sobre o percurso metodológico

Apesar da maior parte das informações terem se baseado nos processos judiciais pesquisados, o levantamento histórico e social não foi o cerne deste trabalho, mas conhecer o funcionamento do órgão e sua relação com a sociedade foi condição imprescindível para atingir o fim delineado pelo projeto, que foi propor um percurso metodológico para a recuperação da informação de um acervo arquivístico, em particular da documentação judicial produzida pelo Juízo de Órfãos do Recife, apontando as metodologias desenvolvidas pela Tipologia Documental e pela Análise Documentária, especificamente, o vocabulário controlado. Como produto final, a proposta foi de oferecer um índice para servir de indicador para futuros trabalhos de produção de instrumentos de busca.

O resultado do tratamento arquivístico, realizado pelo Memorial da Justiça sobre a documentação objeto deste trabalho está disponível em um guia de fundos que contém informações gerais e superficiais sobre o acervo, dividindo a documentação em área cível e criminal, por período cronológico e, no caso da comarca do Recife, por tipo de ação e recurso. Essa identificação está disponível no *site* do TJPE³⁵, bem como em fichas em suporte papel, na própria instituição. Apesar de falhas haverem sido constatadas na classificação desse instrumento de pesquisa, sem ele o caminho para realizar este trabalho seria mais árduo.

A seleção do acervo do Juízo de Órfãos para a aplicação dessa experiência foi definida a partir do conhecimento de dois trabalhos sobre o mesmo Juízo, mas de outras comarcas, tendo como base a documentação de São Paulo (RODRIGUEZ, 2010) e o outro de Pindamonhangaba (ARQUIVO HISTÓRICO MUNICIPAL, 19--). O primeiro, já em forma de tese de doutorado, da área de História Social, apresenta resultados relevantes para o desenvolvimento de estudos sobre Tipologia Documental. O outro, do Arquivo Municipal de Pindamonhangaba, em forma de inventário, traz informações organizadas sobre aquele Juízo, constando tipos de documentos e suas datas, até mesmo com resumo, em alguns casos.

Ao incluir o manuseio das fichas de identificação disponíveis no Memorial para selecionar a documentação para a pesquisa, percebeu-se a necessidade de um aprofundamento acerca das funções e atividades do Juízo de Órfãos, a fim de definir que tipo de ação e recurso deveria ser de sua competência, já que o fichário continha informações sobre a data, o conteúdo da caixa de processos cíveis e o tipo de processo. Depois desse reconhecimento restaram 20 (vinte) caixas para análise. Ao iniciar a leitura dos documentos, perceberam-se

³⁵ Disponível em: <<http://www.tjpe.jus.br/web/memorial-da-justica/guia-de-fundos>>.

impropriedades na classificação descrita e concluiu-se ser necessário reavaliar todas as caixas de processos cíveis referentes ao recorte cronológico, totalizando 44 (quarenta e quatro) caixas. Com isso o quantitativo de caixas duplicou comparando-se a previsão inicial.

Na leitura dos documentos, magistrados, curadores, escrivães, procuradores fiscais e até mesmo advogados apareciam de forma repetida nos processos. Suas caligrafias, apesar de seguirem uma grafia antiga e manuscrita, tornaram-se familiares, facilitando a identificação dos elementos importantes para a coleta.

Elaborou-se uma ficha de identificação com base na Tipologia Documental, estruturada nos elementos descritivos do documento. Para definir os procedimentos para o preenchimento da ficha de coleta de conteúdos, conforme os conceitos da Tipologia Documental, utilizando para tal os elementos intrínsecos como descritores, necessitou-se explorar sistematicamente os processos judiciais, pois na literatura não se identificou a descrição do mesmo tipo documental.

Sabe-se que o processo jurídico é um dossiê composto itens documentais, e que cada um tem papel importante para definir de que trata a ação ou recurso, ou melhor, qual o tipo de processo judicial. Isso decorre da lei que regulamenta uma ação ou recurso. Cita-se como exemplo o caso de um sequestro de bens, em que deverá haver um auto de sequestro a ser executado durante o processo, ou mesmo no caso de um inventário, em que é comum haver um auto de avaliação de bens realizado por peritos.

Após conhecer os tipos documentais manuseando a documentação e testando a ficha para coleta de informações, é que se determinou uma forma padronizada de preenchimento, a partir dos conteúdos apresentados em documentos essenciais de cada um dos processos. Nesta fase da pesquisa percebeu-se que as informações imprescindíveis para alcançar o objetivo do trabalho constavam em determinadas partes do processo, sendo elas:

- a) na capa, como o nome do órgão produtor, da ação ou recurso, a data da autuação;
- b) na petição inicial, ofício ou outro meio em que as partes requeriam em juízo, em que foram encontrados os pedidos daqueles que procuravam o Juízo de Órfãos;
- c) nas decisões do magistrado, como despachos, sentenças e acórdãos, onde se encontravam os direcionamentos dos pedidos e as soluções dos litígios;
- d) nos pareceres de curadores, procuradores fiscais, peritos, onde estavam demonstradas as participações de outros interessados na demanda;
- e) em alguns documentos essenciais, escritos pelos escrivães, tais como os autos de avaliação de bens, autos de arrecadação de bens, autos de sequestro de bens, em que se viu a tramitação dos feitos e os documentos próprios de cada ação e recurso, em particular.

Considerando o processo como uma sequência de atos, conclui-se que o assunto tratado em cada um de seus tipos documentais pode ser identificado pautando-se no conhecimento da sua macroestrutura, conforme descrição das partes selecionadas como detentoras das informações suficientes para fornecer os conteúdos relacionados a cada tipo de ação e recurso específico.

Nesta etapa do trabalho buscou-se um instrumento auxiliar ao processo de recuperação da informação por meio de um recurso que fornecesse ao pesquisador uma lista padronizada de tipos documentais, na qual a indexação dos processos pudesse ser feita a partir de uma lista atualizada de termos. Logo, definiu-se como termos atualizados aqueles definidos pelo Conselho Nacional de Justiça (CNJ) como pontos de acesso para descrição dos processos judiciais de todo o Brasil, existentes desde 2009: as tabelas processuais unificadas de classe e assunto.

Essa escolha é justificada por serem essas tabelas adotadas no cadastro de todos os processos da justiça brasileira, além de ser um parâmetro reconhecido em todos os tribunais e por quem necessita da Justiça nos dias atuais.

Não foi difícil relacionar os tipos documentais do século XIX com os processos comuns ao século XXI, pois os tipos de ação e recurso mudaram pouco no que diz respeito às atribuições do Juízo de Órfãos, apesar do órgão não mais existir. Somente no caso relacionado aos bens ou coisas vagas é que se exigiu a criação de um novo termo, por inexistir na tabela processual unificada de assunto do CNJ. Com relação ao tipo de ação ou recurso, só o *libelo* e a *carta de dia de aparecer* não mais existem. Essas foram mantidas com a nomenclatura usada nos oitocentos.

Importante comentar que os escrivães que serviam à Justiça na comarca do Recife, em Pernambuco, durante o Império, sabiam perfeitamente do que se tratava cada feito, pois a denominação dada e descrita por eles nas capas do processo eram descrições em total sintonia com a lei e o uso daquela época.

Saliente-se que há outro órgão, este privado, que também guarda acervo da mesma comarca estudada. O Instituto Arqueológico, Histórico e Geográfico Pernambucano tem sob sua guarda documentos relacionados à Justiça pernambucana, do período imperial. Contudo, devido à falta de instrumentos de busca existentes no local, sobre o acervo, bem como o acesso restrito a pesquisadores, não foi possível realizar um levantamento prévio da documentação. É provável que outros tipos de ação e recursos não apresentados aqui estejam presentes entre os documentos da comarca do Recife, sob a guarda do Instituto. Entretanto, somente um trabalho futuro poderá confirmar tal suspeita.

4 CONSIDERAÇÕES FINAIS

Percebeu-se a necessidade de congregar distintas metodologias, conforme demonstrado neste trabalho, que se valeu da Arquivologia, da Tipologia Documental e da Análise Documentária (especialmente do vocabulário controlado) para melhor recuperar as informações da documentação arquivística de caráter permanente do Memorial da Justiça (PE). O acervo objeto de análise, composto por processos judiciais, ou seja, com conteúdo vinculado a uma área específica, o Direito, formado com base nesta área, fez gerar a necessidade do aporte de conhecimentos sobre o Direito e da legislação vigente no período em que os documentos foram produzidos.

Depois de alcançado o objetivo deste trabalho, fez-se uma comparação com outras pesquisas similares sobre outros juízos de órfãos, mesmo que de origem diversa. No caso dos tipos de ações listados na tese de Rodriguez (2010) como produção do Juízo de Órfãos de São Paulo, algumas não foram apresentadas como parte do acervo do Memorial. Podem ser citadas, neste caso, a *ação de dez dias* ou *decendial* e a *ação de liberdade* – estando, esta última, presente tanto no trabalho de Rodriguez (2010), quanto no do Arquivo Municipal de Pindamonhangaba (19--).

Para se construir um instrumento de busca completo seria preciso identificar todas as informações contidas na totalidade do acervo de um fundo ou de uma de suas subdivisões. Também seria importante verificar as relações daquele órgão com os outros fundos ou subdivisões, tais como, no caso do Juízo de Órfãos, com o Juízo do Cível e o Juízo da Provedoria, que, por vezes, disputavam competência para julgar processos com aquele Juízo ou mesmo requeriam informações sobre matérias comuns. Mas, para confirmar tal afirmação, seria preciso mais tempo e necessário outro projeto de pesquisa. Fica aqui registrada a sugestão, para que essa proposta possa ser desenvolvida em um novo trabalho.

REFERÊNCIAS

- ARQUIVO HISTÓRICO MUNICIPAL (Pindamonhangaba). **Catálogo Geral Juízo de Órfãos (1701 a 1888)**. São Paulo: Prefeitura Municipal de Pindamonhangaba, 19--. 76p. Disponível em:
<http://www.pindamonhangaba.sp.gov.br/arqhistorico/CatalogosGerais2010/Cat%C3%A1logo-Ju%C3%ADzo%20de%20%C3%93rf%C3%A3os_2011.pdf>. Acesso em: 18 jan. 2013
- BELLOTTO, Heloísa Liberalli. Como fazer análise diplomática e análise tipológica de documento de arquivo. São Paulo: Arquivo do Estado, 2002.
- BELLOTTO, Heloísa Liberalli. **Diplomática e tipologia documental em arquivos**. 2. ed. Brasília, DF : Briquet de Lemos, 2008.
- BEVILÁQUA, Clóvis. **Direito de família**. Recife: Livraria Contemporânea, 1908.

CARDOZO, José Carlos da Silva. Na fronteira da família: entre a lei e a moral. **Em tempo de histórias**, Brasília, n. 17, 2010. Disponível em: <<http://seer.bce.unb.br/index.php/emtempos/article/viewArticle/2877>>. Acesso em: 02 fev. 2012.

CARUCCI, Paola. **Il documento contemporaneo**: diplomatica e criteri di edizione. Roma: La nuova Italia Scientifica, 1987.

CARVALHO, José Pereira de. **Primeiras linhas sobre o processo orfanológico**. Rio de Janeiro: A. A. da Cruz Coutinho, 1879.

CERNE, João Baptista Guimarães. **Ordenações em vigor**: estudo sobre o Código Philippino na nossa atualidade. BDJur, Brasília, DF, 26 mar. 2009. Disponível em: <http://bdjur.stj.gov.br/xmlui/bitstream/handle/2011/20367/Ordена%с3%a7%с3%b5es_vigor.pdf?sequence=6>. Acesso em: 29 mar. 2012.

CUNHA, Mônica Maria de Pádua Souto da. **A análise tipológica como subsídio para construção de vocabulário controlado**: o caso do Juízo de órfãos do Recife (1824-1889) / Mônica Maria de Pádua Souto da Cunha. – Recife: O autor, 2013. 235f. Originalmente apresentada como dissertação de mestrado, Departamento de Ciência da Informação, Universidade Federal de Pernambuco.

CUNHA, Mônica Maria de Pádua Souto da; ALMEIDA, Alessandra Bárbara Santos. A informação arquivística do Judiciário e o seu valor permanente: o usuário no contexto da avaliação documental. In: CONGRESSO BRASILEIRO DO PODER JUDICIÁRIO, 2., 2012, Brasília, DF. **Anais...** Brasília, DF: [s.n.], 2012.

DURANTI, Luciana. **Diplomática**: nuevos usos para una antigua ciencia. Carmona: Asociación de Archiveros de Andalucía, 1995.

GUIMARÃES, José Augusto Chaves. Formas da informação jurídica: uma contribuição para sua abordagem temática. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 26, n. 1/2, p. 41-54, jan./jun., 1993.

HERRERA, Antonia Heredia. **Archivística general teoria y practica**. Sevilha: Diputación Provincial de Sevilha, 1988.

HERRERA, Antonia Heredia. **El principio de procedencia y los otros principios de La archivística**. São Paulo: Associação de Arquivistas de São Paulo, 2003.

KOBASHI, Nair Yumiko. **Vocabulário controlado**: estrutura e utilização. Brasília: Enap, 2008. 4 f. Disponível em: <http://www2.enap.gov.br/rede_escolas/arquivos/vocabulario_controlado.pdf>. Acesso em: 09 abr. 2012.

LANCASTER, Frederick Wilfrid. **Indexação e resumos: teoria e prática**. 2. ed. Brasília: Briquet de Lemos, 2004.

LYRA, Augusto. Tavares. Organização política e administrativa do Brasil (Colônia, Império e República). São Paulo: Companhia Editora Nacional, 1941.

MALHEIRO, Agostinho Marques Perdigão. **A escravidão no Brasil**: ensaio histórico, jurídico-social. Rio de Janeiro: Typografia nacional, 1866. v. 1.

NEVES, Lúcia Maria Bastos Pereira das; MACHADO, Humberto Fernandes. **O império do Brasil**. Rio de Janeiro: Nova Fronteira, 1999.

RODRIGUES, Ana Célia. **Diplomática contemporânea como fundamento metodológico da identificação de tipologia documental em arquivos**. São Paulo: O autor: 2008. 258 f. Tese (Doutorado em História social)-Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, São Paulo, 2008. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/8/8138/tde-27112008->. Acesso em: 15 Jan. 2013.

RODRIGUEZ, Sonia Maria Troitiño. **O Juízo de Órfãos de São Paulo**: caracterização de tipos documentais (séc. XVI-XX). São Paulo: o autor: 2010. 283 f. Tese (Doutorado em História Social)-Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, São Paulo, 2010. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/8/8138/tde-27042010-103207/pt-br.php>>. Acesso em: 15 Jan. 2013.

RONDINELLI, Rosely Curi. **Gerenciamento arquivístico de documentos eletrônicos**: uma abordagem teórica da diplomática arquivística contemporânea. 3. ed. Rio de Janeiro: Ed. FGV, 2005.

SCHELLENBERG, Theodore R. **Arquivos modernos**: princípios e técnicas. 6. ed. Rio de Janeiro: Ed. FGV, 2006.

SILVEIRA, Luiz de Souza da. **Anotações à Lei n. 2040, de 28 de Setembro de 1871**: seguidas de todas as leis e decretos relativos a escravidão, tráfico de africanos, locação de serviços, avisos e decisões dos tribunais, e de um formulário de todas as acções. Maranhão: Frias, 1876.

SMIT, Johanna Wilhelmina. KOBASHI, Nair Yumiko. **Como elaborar vocabulário controlado para aplicação em arquivos**. São Paulo: Arquivo do Estado, Imprensa Oficial do Estado de São Paulo, 2003.

ENTRE O USUÁRIO E A INFORMAÇÃO CORPORATIVA: A BUSCA POR FACETAS

BETWEEN USERS AND ENTERPRISE INFORMATION: FACETS OF THE SEARCH

Leonardo Lacerda Alves

Resumo: Este trabalho objetiva responder quais características da informação corporativa manifestam-se tanto em documentos quanto em expressões de busca de seus usuários. A metodologia é descritiva, aplicada e experimental. A coleção de documentos adotada foi produzida a partir de um repositório corporativo em uso dentro de uma empresa real. Termos foram extraídos do conteúdo de documentos e de expressões de busca de usuários. Os termos foram categorizados e foi aplicada a análise facetada para descobrir facetas comuns, úteis tanto para o conteúdo dos documentos quanto para o conteúdo das expressões de busca. Oito categorias de assuntos são tão frequentes em documentos quanto nas expressões de busca, apresentando alto e positivo coeficiente de correlação de Spearman. A análise facetada constituiu um instrumento útil para a comparação de assuntos presentes em documentos e nas expressões de busca sem que dados sensíveis da empresa fossem expostos. O método empregado neste trabalho também apresenta potencial para reconhecer características locais e gerais da informação corporativa que deem suporte ao desenvolvimento de sistemas automáticos de recuperação de informação corporativa mais eficazes.

Palavras-chave: Análise de domínio. Informação corporativa. Classificação facetada.

Abstract: The paper endeavours to expose the main characteristics of the enterprise information, using documents and queries the employees use to search the documents. A descriptive, applied, and experimental methodology was devised to investigate the enterprise information characteristics by selecting terms from the documents and the queries, by applying facet analysis on the enterprise collection, and comparing the resulting categories and facets from the documents with those from the queries. The subject distributions were compared using the Spearman's rank correlation. Eight categories were determined by identifying the subjects and their distribution present strong positive coefficient. The adopted method can recognise local and general characteristics of enterprise information to support the development of more efficient enterprise information retrieval systems. The facet analysis was useful for comparing the subjects in documents and user queries with low potential to expose confidential data. Thus, the method can guide and stimulate future work and other companies can become more willing to take part in a research study.

Keywords: Domain analysis. Enterprise information. Faceted classification.

1 INTRODUÇÃO

Empresas detêm um volume expressivo de dados e, para que trabalho e decisões sejam realizados com eficiência, seus atores sociais dependem de uma eficaz recuperação de informação. Como o volume de dados é crescente, a adoção de processos automáticos e semiautomáticos de organização da informação torna-se essencial. Porém, uma parte expressiva da informação corporativa possui utilidade limitada a pequenos intervalos de tempo e para poucos usuários. Além disso, diferentes contextos de trabalho fazem com que a informação apresente atributos diversificados, mesmo em uma só empresa. Caracterizar essa

diversidade é fundamental para dar suporte a melhores serviços de informação.

Identificar características da informação corporativa não é trivial. Dentro de uma empresa há diversidade de necessidades de informação, de fluxos de trabalho, de competências humanas e de recursos tecnológicos. Se em uma única empresa isso é verdade, a tarefa de conhecer o domínio corporativo é ainda mais desafiadora, situação na qual empresas de diferentes portes, atividades econômicas, localidades, estruturas organizacionais e idiomas compartilham de necessidades informacionais aparentemente compatíveis apenas em sua complexidade. Esse é apenas o primeiro grande obstáculo para o desenvolvimento de sistemas padronizados para a indexação e recuperação automáticas de informação corporativa, com custos humanos e tecnológicos menores que o processamento intelectual ou que o desenvolvimento de um sistema diferente para cada unidade organizacional.

Para avançar na direção de um sistema de recuperação de informação corporativa, dois requisitos concomitantes são essenciais. Primeiro, é preciso que sejam conhecidas as características mais comuns pelas quais os usuários buscam a informação dentro da empresa. Segundo, é preciso que sejam conhecidas as características mais comuns da informação em todo o domínio corporativo. Ambos os requisitos dependem de uma profunda análise do domínio corporativo e ajudam no desenvolvimento de sistemas de recuperação de informação corporativos mais eficazes, enquanto o primeiro vai além: reconhece as necessidades específicas e favorece uma fina sintonia dos processos de indexação e recuperação ao contexto particular de cada empresa. Esta pesquisa concentra-se no primeiro requisito.

Este trabalho concentra-se na seguinte pergunta: Quais características da informação corporativa manifestam-se tanto em documentos quanto em expressões de busca de seus usuários? Assim, esta pesquisa objetiva propor um conjunto de facetas que auxiliem na representação do conteúdo organizacional e que possam ser utilizadas como expressões de busca, ou consultas, de seus usuários, fazendo uso de um método que possa ser facilmente repetido em trabalhos futuros. Os usuários desempenham os papéis de autores e receptores das mensagens corporativas, sendo esperado que eles reconheçam, mesmo que inconscientemente, as características mais importantes da informação. Nas empresas, documentos são os veículos mais comuns de mensagens entre os diversos atores sociais. Para realizar trabalho e tomar decisões, esses atores frequentemente recuperam documentos, de sua autoria ou não, fazendo uso de termos em linguagem natural e partindo até mesmo de conselhos e sugestões de colegas de trabalho que passaram anteriormente por aquela necessidade informacional (CHOO *et al.*, 2008; WU *et al.*, 2014). Apesar de útil e desejável, um vocabulário controlado nem sempre está disponível em unidades organizacionais. Assim,

os usuários tendem a levar atitudes e estratégias de busca da *Web* para o contexto corporativo, o que não funciona plenamente e produz frustração e ineficiência (LIU *et al.*, 2012).

Partindo do pressuposto que há características comuns entre documentos corporativos e expressões de busca de usuários corporativos, em diferentes unidades organizacionais e empresas, a análise de domínio que pode representá-las enfrenta três limitações. A primeira refere-se à resistência de empresas em disponibilizar uma amostra significativa da sua informação, muitas vezes estratégica e sensível, a qual expõe funcionários, parceiros e clientes (BAILEY *et al.*, 2007). A segunda limitação considera a hipótese de que o domínio corporativo não é suficientemente homogêneo para que um único modelo represente todo o domínio (HALEVY *et al.*, 2005). A terceira refere-se à limitação dos métodos quantitativos adotados para reconhecer características da informação principalmente em áreas externas à Ciência da Informação. Tal reconhecimento requer estudos qualitativos aprofundados, históricos e racionais sobre a natureza, propósitos e uso da informação corporativa (HJØRLAND, 2002). Não é trivial contornar as limitações citadas, mas um caminho para reduzi-las é a escolha de procedimentos metodológicos que permitam a adequada avaliação de informação e exponham menos os dados. A análise facetada parece ser um caminho.

A origem da classificação facetada se deu com estudos de Ranganathan (1967) e teve continuidade nos estudos de outros autores da área de Biblioteconomia e Ciência da Informação, apesar do conceito estar em uso crescente também em outras áreas com algumas variações (BROUGHTON, 2006; VICKERY, 2008). As facetas identificadas nesta pesquisa constituem características da informação de uma única empresa, mas favorecem a análise do domínio corporativo na medida em que o método pode ser repetido em outras organizações, usando documentos de diferentes tipos, propósitos e idiomas, e usando consultas em contextos diversos de trabalho. As facetas evidenciadas neste trabalho e em trabalhos futuros poderão ser comparadas sem que os dados sejam expostos, algo que representa uma grande vantagem para as organizações, para o campo de estudo e para os profissionais de informação envolvidos no desenvolvimento de sistemas de recuperação de informação corporativa.

Na seção seguinte são apresentados os fundamentos teóricos e a revisão da literatura sobre análise de domínio, análise facetada e classificação da informação corporativa. Os procedimentos metodológicos são detalhados na seção três, seguida pela apresentação dos resultados e sua discussão, nas seções quatro e cinco. Finalmente, a seção seis apresenta conclusões e limitações deste estudo e aponta algumas direções de trabalhos futuros.

2 FUNDAMENTOS TEÓRICOS E REVISÃO DA LITERATURA

Na teoria da análise facetada de Ranganathan (1967) faceta é um conceito central, embora apresente interpretações variadas na literatura (BROUGHTON, 2006). Porém, faceta sempre é vista como uma característica permanente e relevante de uma entidade, algumas vezes reconhecida por meio da análise facetada. Análise facetada é uma técnica através da qual facetas são descobertas em um domínio e normalmente formam uma classificação facetada (LA BARRE, 2010). A abordagem facetada constitui uma estratégia classificatória pela qual tenta-se responder como uma entidade pode ser descrita, ao invés de tentar responder onde a entidade seria posicionada dentro de uma estrutura classificatória definida previamente (RANGANATHAN, 1967). Assim, variadas perspectivas e incertezas sobre necessidades de usuários são consideradas, dentro de um processo classificatório incremental e iterativo, tornando seu produto útil para mais de um problema (HJØRLAND, 2002).

Por meio de facetas significativas, o domínio corporativo pode ser melhor organizado, a organização da informação pode ser aperfeiçoada e sistemas de recuperação de informação mais eficientes são possíveis (ALBRECHTSEN, 1993). Os requisitos para a análise de todo um domínio e para a análise da informação de uma única empresa são diferentes (GOPINATH, 2002). Enquanto o domínio requer uma compreensão mais geral e consensual, a informação de uma instituição admite perspectivas baseadas em maior empirismo e pragmatismo, sendo útil apenas em um contexto mais limitado (HJØRLAND, 2012).

Mesmo havendo muitas características relevantes no domínio, o conjunto de facetas consideradas restringe-se às mais importantes por razões econômicas e práticas (RANGANATHAN, 1967), algo que aumenta a importância e o rigor da análise de um domínio para incluir características preferencialmente reconhecidas por todos os atores do domínio e em qualquer tipo de documento (LYKKE-NIELSEN, 2011). Por outro lado, as características locais da informação corporativa, a terminologia específica de uma atividade econômica, as necessidades individuais de cada unidade organizacional e o contexto geográfico e histórico da organização são representadas da mesma forma como aquelas mais gerais do domínio (DOLBY *et al.*, 2009; GIESS *et al.*, 2008; WILD *et al.*, 2009), mesmo não sendo compartilhadas por todas as empresas. Também, os usuários da informação estão interessados tanto em dados explícitos presentes no conteúdo de documentos (GARDIN, 1973) quanto em informação implícita e conhecimento informal distribuído pelos sistemas corporativos (ALWIS; HIGGINS, 2001; CHOO *et al.*, 2008; MARCELLA; ILLINGWORTH, 2012; NUNES *et al.*, 2006; O'FARRILL, 2010), o que representa um desafio adicional para a representação do domínio. Portanto, o domínio, no nível mais global ou mais local, é

dinâmico e sua representação precisa estar preparada para atualizações contínuas.

O método facetado mostra-se útil para o domínio corporativo porque favorece uma análise incremental de um domínio usando uma base teórica sólida a partir da Biblioteconomia e da Ciência da Informação (GARFIELD, 1984). Embora sejam mais populares na sua área de origem e para a classificação bibliográfica, as abordagens facetadas estão se popularizando em outras áreas (LA BARRE, 2010). São exemplos não exaustivos os documentos de engenharia (GIESS *et al.*, 2008; WILD *et al.*, 2009), *sites* corporativos (WANG *et al.*, 2008), documentos *Web* (HONG, 2006; VICKERY, 2008) e catálogos *Web* (SACCO, 2007). Como a diversidade de aplicações e disciplinas tem aumentado, surgem também visões alternativas e um grande número de trabalhos sem que todos os princípios originais de Ranganathan sejam respeitados (SPITERI, 1998; WILD *et al.*, 2009).

A análise facetada da informação corporativa como uma técnica de análise de um domínio deve começar por definir as facetas mais comuns e frequentes à maioria das empresas. As sugestões mais tradicionais partem das cinco categorias consideradas fundamentais por Ranganathan (1967), personalidade, matéria, energia, espaço e tempo, e as treze recomendadas pelo *Classification Research Group*: coisa, tipo, parte, propriedade, material, processo, operação, agente, paciente, produto, subproduto, espaço e tempo. Elas supostamente têm o potencial de representar o conhecimento de qualquer domínio, sendo usadas como ponto de partida para a análise facetada (BROUGHTON, 2006).

Portanto, nenhuma das categorias sugeridas anteriormente é característica exclusiva da informação corporativa. Entretanto, elas são potenciais candidatas por pertencerem às entidades corporativas e não-corporativas presentes nos documentos produzidos pelas empresas. Na análise facetada empreendida neste trabalho, é um requisito que as categorias sejam boas candidatas a atributos permanentes que organizem e relacionem as diversas entidades presentes na comunicação empresarial ao longo do tempo e do espaço (LA BARRE, 2010). Espera-se, também, que tais candidatas constituam um modelo de longo prazo para a organização da informação corporativa, requerendo poucas revisões e favorecendo o desenvolvimento de sistemas de recuperação de informação corporativa, suficientemente flexíveis, para integrar toda e qualquer unidade organizacional da empresa (VAKKARI; JÄRVELIN, 2005). Dessa forma, as categorias identificadas na análise de um domínio seriam mapeadas em facetas de uma classificação facetada para representar a informação corporativa.

Diferentes estratégias têm sido adotadas para o reconhecimento de facetas relevantes para representar informação corporativa, como construção de ontologias, análises de redes sociais e até mesmo adoção da *Web* e dados ligados para reconhecimento e desambiguação de

facetadas sociais e espaciais (DAVIS JR *et al.*, 2011; BUKOWSKA *et al.*, 2012). Após a adoção dessas estratégias, bancos de dados, ontologias e serviços estruturados têm sido usados para dar suporte ao processamento especial de dados não estruturados, inclusive dados complementares de terceiros (DAVIS JR; ALVES, 2007). De acordo com Halevy *et al.* (2005), a informação corporativa normalmente não é estruturada ou é semiestruturada, apesar de haver um grande volume de dados estruturados em bancos de dados que não conta com esquemas semânticos associados. Esse é o principal desafio para o processamento automático da informação corporativa, algo não distante da realidade do contexto de documentos *Web* (HONG, 2006; OVERELL, 2009; VICKERY, 2008).

Em ambos os contextos, as facetadas de entidades são usadas como suporte para a busca automática (LIU *et al.*, 2012), *browsing* (OREN *et al.*, 2006; WANG *et al.*, 2008) e ordenação de resultados (GIRGESOHN *et al.*, 2010), na *Web* e fora dela. Além do suporte a vários processos, as pesquisas têm apontado a utilidade do tratamento de várias facetadas combinadas (CARDOSO; SANTOS, 2008; SOLOMON, 2002), ao invés do tratamento individualizado de facetadas, quando são adotadas apenas as espaciais (AUER *et al.*, 2009), temporais (RULA *et al.*, 2012) ou sociais (ABEL *et al.*, 2011; GUY *et al.*, 2012). Em conjunto, facetadas diferentes se complementam e suportam as necessidades dos usuários com maior precisão (BROUGHTON, 2006). Isso indica o porquê de usuários escolherem usar, por exemplo, facetadas espaciais e temáticas em motores de busca (BORGES *et al.*, 2007; JONES *et al.*, 2008). Se usuários de informação corporativa têm o mesmo comportamento ainda é uma questão em aberto, algo que esta pesquisa começa a responder.

3 PROCEDIMENTOS METODOLÓGICOS

Este trabalho é descritivo, aplicado e experimental. Os métodos adotados têm a função de caracterizar os documentos de uma coleção corporativa e as consultas de usuários ao buscar por um conjunto de documentos.

A coleção de documentos adotada é um conjunto de documentos escritos em língua portuguesa, pertencente a um campus do Centro Federal de Educação Tecnológica de Minas Gerais (CEFETMG). São 1305 documentos que incluem principalmente e-mails, páginas *Web*, relatórios técnicos, notícias internas e externas, e pautas e atas de reunião. Os documentos foram criados entre os anos de 2007 e 2013 e têm sido acessados por seus usuários durante o período de 2012 a 2013.

De uma população de sete coordenadores, cinco se prontificaram a participar da pesquisa. Eles são familiarizados com o uso dos documentos da coleção e foram solicitados a

propor consultas que levassem para os mesmos dez documentos da coleção de documentos, pré-selecionados aleatoriamente.

Os cinco participantes foram coordenadores durante o período de 2007 a 2014 ou são coordenadores atualmente. Todos são professores do sexo masculino, sendo que dois deles pertencem à área de Engenharia Elétrica e três à área de Ciência da Computação. Dois concluíram o doutorado e três concluíram o mestrado antes de exercerem o cargo.

A proposta de participação na pesquisa foi enviada através do e-mail profissional. A mensagem do e-mail trazia a proposta da pesquisa, seu formato e aspectos éticos, as instruções, exemplos de consulta e o endereço do formulário on-line através do qual as consultas deveriam ser registradas. Em anexo a esse e-mail foram incluídos onze documentos, um usado para algumas consultas de exemplo, que simplesmente ilustravam o que constituía uma expressão de consulta, e dez que compõem a amostra. Os participantes registraram as consultas sem usar sistemas ou repositórios da empresa, tendo acesso irrestrito à Internet. Responderam a pesquisa quando estavam em casa, sem prazo limite de participação.

A amostra foi produzida aleatoriamente por não haver qualquer evidência de vício na coleção, supostamente composta apenas por documentos conhecidos e úteis para a unidade organizacional. O número de documentos, por outro lado, foi decidido considerando a utilidade para a pesquisa e a viabilidade para os participantes. Era esperado que um número reduzido de documentos permitisse aos participantes realizar as consultas de uma única vez, sem que a participação fosse interrompida por obrigações profissionais ou domésticas.

A consulta proposta deveria ser escrita em linguagem natural, pois não há vocabulário controlado para a coleção de documentos. Porém, como cada documento de interesse era consultado antes da elaboração de sua consulta, era esperado que o vocabulário presente no documento fosse adotado por todos os participantes e influenciasse as expressões de busca.

Para reduzir algumas das limitações do método e permitir pesquisas que possam ser comparadas com a atual, os seguintes requisitos foram definidos: a amostra de documentos é aleatória; uma cópia dos documentos foi entregue aos participantes, não requerendo acesso ao repositório da empresa; os participantes foram instruídos a não usarem qualquer recurso além dos documentos providos; não foi fornecida orientação sobre como se fazer uma busca eficiente ou quais e quantos termos deveriam ser usados na consulta; e os participantes foram orientados a consumirem o tempo de dez minutos para elaborar todas as consultas.

Um dos participantes é autor deste trabalho e foi o responsável pela categorização dos documentos da coleção. Portanto, como o objetivo deste trabalho é comparar características dos documentos e das consultas dos usuários, seu formulário foi descartado e os resultados da

pesquisa não usam as consultas propostas pelo autor-participante. Sua participação serviu apenas para testar o formulário e estimar o tempo de resposta. Na seção seguinte são apresentados os procedimentos para a identificação de assuntos mais significativos na coleção e nas consultas, e para a categorização dos seus termos.

3.1 Análise de assunto e análise facetada

Foi feita a análise de assunto e posteriormente aplicada a análise facetada para a determinação das facetas em todos os 1305 documentos e nas consultas de quatro participantes, totalizando 40 consultas.

Os documentos foram indexados e categorizados por um único pesquisador. A indexação produziu uma lista de termos em linguagem natural que pertencem à coleção, tendo sido categorizados ao final da indexação. São as categorias e dois níveis de subcategorias os objetos de estudo neste trabalho, sendo que nem sempre os três níveis foram mobilizados. Um exemplo de termo em que apenas um nível foi mobilizado é ‘denúncia’, categorizado como Comunicação. O termo ‘departamento’ mobiliza dois níveis, a categoria Instituição e a subcategoria Unidade. O termo ‘diretor’, por sua vez, mobiliza os três níveis, a categoria Pessoal, a subcategoria Profissional e o terceiro nível Papel.

As categorias resultantes da coleção de documentos são consistentes em função do emprego de um único indexador. Indexadores adicionais para a coleção de documentos são dispensáveis para o objetivo deste trabalho. Por outro lado, a consistência entre as categorias da coleção de documentos e as categorias das consultas deve ser medida, uma vez que os termos da coleção e os termos das consultas foram produzidos por participantes diferentes, constituindo objeto de estudo deste trabalho.

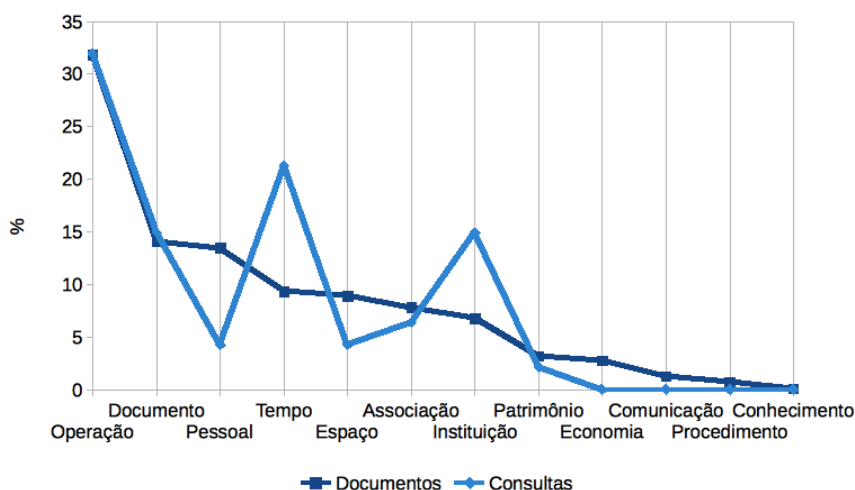
As categorias e subcategorias resultantes foram usadas para comparar as características dos assuntos que compõem o conteúdo dos documentos e o conteúdo das consultas dos usuários corporativos, respondendo ao problema da pesquisa e atendendo a seus objetivos. Para avaliar o padrão da distribuição de assuntos entre coleção e consultas foi adotado o coeficiente de correlação de Spearman. Ele se mostrou adequado para a comparação, além de não requerer uma análise estatística excessivamente complexa. Portanto, o coeficiente de Spearman também é útil para que os resultados presentes e de trabalhos futuros sejam comparados sem que os dados de empresas sejam expostos.

4 RESULTADOS

O principal objetivo deste trabalho foi identificar as categorias ou características comuns a documentos corporativos e consultas de usuários enquanto buscam por documentos

pertinentes em seu contexto de trabalho. O processamento integral da coleção de documentos consumiu 29 horas e 20 minutos de um único pesquisador. O tempo médio de processamento de cada documento da coleção foi de 1 minuto e 21 segundos. Adicionalmente, dez consultas comuns foram produzidas por quatro participantes da empresa investigada, totalizando 40 consultas. O tempo médio de produção de cada consulta é igual a 1 minuto. A categorização dos termos presentes nas consultas foi realizada pelo mesmo indexador dos documentos, o que consumiu um tempo adicional de 2 horas e 22 minutos.

FIGURA 1 – Distribuição de assuntos entre as categorias

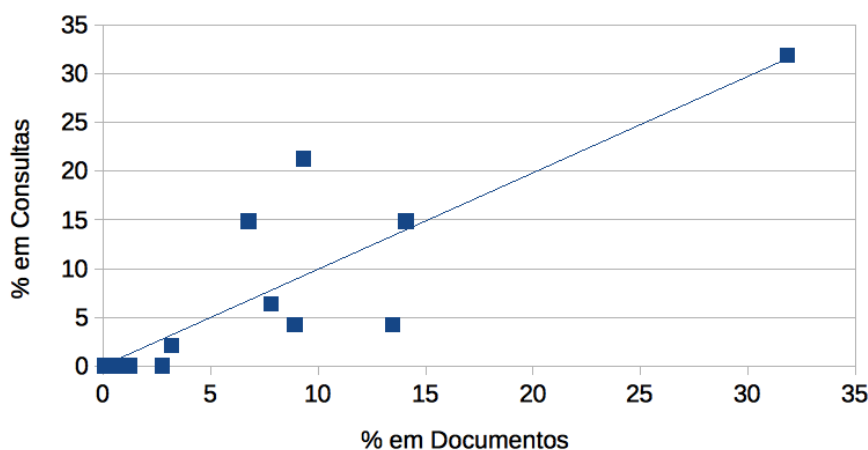


Fonte: o autor.

O indexador identificou 1643 termos/assuntos diferentes em toda a coleção de documentos e 47 termos/assuntos diferentes nas consultas de usuários. Os assuntos identificados durante a indexação dos documentos e na formulação de consultas foram agrupados em 12 categorias definidas no momento em que a categorização acontecia. A figura 1 apresenta as categorias e a proporção de assuntos associados a cada uma delas. As categorias são Pessoal, Associação, Instituição, Documento, Comunicação, Economia, Conhecimento, Operação, Procedimento, Espaço, Patrimônio e Tempo. A proporção de assuntos em cada categoria é compatível entre a coleção de documentos e nas consultas formuladas por usuários da informação corporativa. As categorias que apresentaram maior compatibilidade entre os conjuntos avaliados são Operação, Documento, Associação e Patrimônio, com a frequência de assuntos quase idêntica entre documentos e consultas. As categorias Economia, Comunicação, Procedimento e Conhecimento, embora apresentem

pontos graficamente próximos na figura 1, não apresentaram qualquer assunto nas consultas formuladas e portanto devem ser descartados da análise.

FIGURA 2 – Correlação da distribuição de assuntos entre categorias



Fonte: o autor

A FIGURA 2 ilustra a dispersão das frequências de assuntos em categorias entre documentos e consultas. A distribuição apresenta alta correlação positiva indicada por coeficiente de Spearman: $\rho(10) = 0,860912$, $n = 12$, $p = 0,00032$, usando as 12 categorias.

Por outro lado, nos níveis mais baixos da categorização, houve uma reduzida correlação positiva entre a distribuição de assuntos em facetas e subfacetas entre os documentos e as consultas. Isso sugere a inexistência de correlação estatística na medida em que o nível de classificação torna-se mais preciso.

Especificamente na amostra de consultas, os usuários participantes fizeram uso de um total de 47 termos diferentes para elaborar as consultas que levariam aos dez documentos de interesse. O quadro 1 apresenta os termos adotados pelos usuários para os cinco primeiros documentos da amostra. Para cada documento, os termos usados pelos usuários que pertencem a mesma categoria foram alinhados horizontalmente. Assim, para buscar pelo documento 1, os quatro usuários fizeram uso dos termos “congregação” e “ata” das categorias Operação e Documento. Apenas dois fizeram uso do termo “2009” e um fez uso do termo “dezembro de 2009”, ambos da categoria Tempo. Três deles fizeram uso do termo “timóteo” da categoria Espaço e “cefet” de Instituição. Dois deles fizeram uso do termo “reunião” da categoria Operação. Finalmente, apenas um deles fez uso do termo “campus” da categoria Instituição. O QUADRO 2 faz a mesma apresentação para os últimos cinco documentos.

QUADRO 1 – Termos usados em consulta para busca por documentos de 1 a 5

N. Doc.	Usuário 1	Usuário 2	Usuário 3	Usuário 4
1	congregação ata 2009 timóteo cefet campus reunião	congregação ata dezembro de 2009	congregação ata timóteo cefet	congregação unidade ata 2009 timóteo cefet reunião
2	regime disciplinar discente	regime disciplinar discente	regime disciplinar cefet timóteo	regime disciplinar discente cefetmg
3	palestra edificações timóteo	palestra edificações incêndio	palestra timóteo combate incêndio cefet	palestra curso técnico edificações timóteo fundamentos para elaboração projeto prevenção cefet
4	2010 greve professor publicação externa 2012	greve timóteo professor	greve timóteo cefet diário do aço	timóteo cefet dois meses de paralisação
5	estrutura organizacional timóteo 2014	organograma timóteo	estrutura organizacional timóteo cefet	estrutura organizacional timóteo cefet

Fonte: o autor.

Como pode ser observado nos quadros 1 e 2, os usuários participantes adotaram um número muito diverso de termos. O usuário 1 usou um número médio de 3,6 termos para cada consulta, valor próximo da média de 3,8 termos do usuário 3. O menor valor médio pertenceu ao usuário 2, com 2,6 termos para cada consulta, e o maior valor médio pertenceu ao usuário 4, com 4,6 termos para cada consulta.

As consultas, por sua vez, apresentaram diferenças maiores. A consulta 1 foi formulada com o maior número de termos, 5 em média, e com o maior desvio-padrão 1,825. A consulta 2, por sua vez, foi formulada com o menor número de termos, 1,75 em média. As consultas, como um todo, foram formuladas com um número médio de termos igual 3,625, com desvio-padrão de 1,314.

QUADRO 2 – Termos usados em consulta para busca por documentos de 6 a 10

N. Doc.	Usuário 1	Usuário 2	Usuário 3	Usuário 4
6	requerimento coordenação de programa de estágios	requerimento	requerimento cefet timóteo estágio	requerimento coordenação programa estágio cefet timóteo
7	curso pró-técnico 2009	estágio carta vestibular pró-técnico 2010	vestibular	vestibular curso preparatório
8	publicação externa fórum dos coordenadores ata	fórum ata março de 2013	prefeitura timóteo cefet fórum ata cefet timóteo	prefeitura timóteo cefet fórum coordenadores ata 2013 cefet timóteo
9	sistema recuperação palestra mark	palestra mark	palestra mark cefet timóteo	2 reunião palestra prof mark cefet timóteo
10	2013 realidade aumentada congresso abm	congresso abm expominas	congresso abm cefet timóteo	64 congresso abm cefet timóteo

curso técnico em metalurgia

Fonte: o autor.

5 DISCUSSÕES

A seção anterior apresenta os resultados da categorização de assuntos dos documentos corporativos e das consultas formuladas pelos usuários da empresa pesquisada. Os resultados atendem ao objetivo de explicitar características comuns aos documentos e consultas por meio da distribuição de assuntos entre as categorias identificadas.

A indexação dos documentos da coleção e a identificação de 12 categorias de mais alto nível aconteceu no contexto de uma pesquisa de doutorado em andamento. Portanto, a análise de domínio preliminar aconteceu antes de considerar os usuários e suas consultas como objeto de estudo. A amostra de consultas, apesar de reduzida, é suficiente para determinar a compatibilidade das categorias entre o conjunto de documentos e de consultas. Porém, ao

incluir as consultas de usuários, foram obtidos 47 termos/assuntos, categorizados por meio de apenas 8 das 12 categorias originalmente encontradas nos documentos.

Em metade das 8 categorias, a frequência de assuntos que ocorre na coleção de documentos e nas consultas é quase idêntica. É esse resultado que justifica a adoção do método em pesquisas futuras e merece ser verificado em outras empresas e em outras situações de trabalho da mesma empresa investigada. No entanto, a frequência desigual entre as amostras de algumas categorias, os pontos mais distantes da linha de tendência da figura 2, não significa prontamente que tais categorias sejam menos importantes. A figura 1 mostra as categorias Tempo e Instituição, por exemplo, com mais de 15% dos assuntos informados nas consultas pelos usuários. Claramente, isso indica que o tempo e as unidades organizacionais são mais importantes no momento de consultar, restringindo a informação que se deseja obter, do que no momento de produzir, quando o tempo segue seu curso natural e as unidades organizacionais apresentam um ciclo de vida mais longo.

Pessoal é outra categoria onde a frequência não é compatível. Em uma amostra tão reduzida, os participantes fizeram pouco uso de nomes de pessoas e papéis desempenhados na instituição. E como já citado, fizeram uso intenso de assuntos da categoria Instituição. Aparentemente, esses assuntos impuseram um filtro artificial à informação, pois os usuários reconhecem que seus documentos de interesse constituem uma fração muito pequena da miríade de documentos que sua instituição tem produzido ao longo dos anos. Tais fenômenos não ocorrem entre os assuntos presentes nos documentos. Muitos atores sociais aparecem nas mensagens e muitos são os seus papéis e cargos. Em número muito inferior ao de pessoas, as unidades organizacionais estão sempre presentes, mas se repetem em vários documentos da empresa, minimizando sua participação no conjunto de assuntos presentes nas mensagens.

Por outro lado, quatro categorias não apresentaram assuntos em consultas, a saber: Economia, Comunicação, Procedimento e Conhecimento. Isso se justifica pelo tamanho da amostra limitado a apenas dez documentos de interesse. O acréscimo de mais documentos certamente produziria a ocorrência de assuntos de categorias não mobilizadas.

De qualquer forma, a correlação ilustrada na figura 2, a partir das 12 categorias e não apenas das 8 que apresentaram assuntos nas amostras, confirma a compatibilidade da frequência de assuntos em categorias entre os documentos e as consultas formuladas pelos usuários. As quatro categorias em que não houve assunto entre as consultas também apresentam escassez de assuntos entre os documentos, como se observa na figura 1.

As categorias mais numerosas são Operação, Documento e Pessoal, que juntas correspondem a aproximadamente 50% de todos os assuntos tratados nos documentos e

propostos por usuários nas consultas. Isso explica o porquê de os sistemas de recuperação de informação corporativa atuais serem principalmente baseados em ontologias corporativas mais generalistas, na classificação de documentos por seu tipo ou gênero, e no reconhecimento de pessoas, especialmente especialistas e autoridades da empresa. Falta responder se a outra metade de assuntos tem sido considerada por indexadores intelectuais e automáticos apenas como termos comuns em uma busca em texto completo ou como categorias especiais da informação corporativa.

A quantidade de termos usados nas consultas foi cerca de 1,6 vezes maior que a média encontrada na literatura para usuários da *Web*, sendo que apenas 32,2% das consultas apresentaram três ou mais termos (LIU *et al.*, 2011; SPINK *et al.*, 2001). No entanto, a maior incompatibilidade é a própria definição do conceito de termo, sendo usado na maioria das vezes como sinônimo de qualquer sequência de caracteres alfanuméricos, uma palavra (SPINK *et al.*, 2001). Comparando dessa forma, o número de palavras usadas por participantes é ainda maior, em média 4,7 palavras por consulta, com desvio-padrão de 1,713, duas vezes maior que a média de 2,16 palavras por consulta na *Web*.

Entretanto, considerar “*new york*” como dois termos, ao invés de um, não parece apropriado e muito menos necessário. Entre os 50 pares mais frequentemente encontrados por Spink *et al.* (2001) também estão “*united states*”, “*real estate*”, “*new jersey*”, “*north carolina*”, “*for sale*”, “*university of*”, “*home page*”, “*high school*”, “*university state*” e “*chat rooms*”, todos candidatos a representarem apenas um termo composto por duas ou mais palavras, algo trivial de resolver por meio de um dicionário. Assim, comparar o número de termos deste trabalho com o número de palavras presente na literatura é mais apropriado e nosso número 1,6 vezes maior parece compatível.

Uma hipótese útil refere-se ao comportamento do usuário em motores de busca na *Web* estar sendo transferido para sistemas de recuperação de informação dentro da empresa. Isso pode ser verdade especialmente quando observados usuários sem treinamento para o uso do sistema de recuperação de informação corporativa ou sem formação acadêmica que o especialize a recuperar informação específica.

Por outro lado, a categorização dos termos elaborados nas consultas apresenta valores que diferem de resultados vistos na literatura para *Web*. Borges *et al.* (2007) sugerem que entre 14 e 18% das consultas formuladas na *Web* apresentam termos da categoria Espaço. Em nosso contexto corporativo, os usuários 3 e 4 usaram um termo da categoria Espaço em todas as consultas, sempre em conjunto com um termo da categoria Instituição, com o propósito de ajudar o sistema a recuperar apenas informação da unidade organizacional em que trabalham.

Porém, desconsiderados os usuários 3 e 4, em 50% das consultas pelo menos um dos demais usuários fez uso de um termo da categoria Espaço aparentemente sem esse propósito.

Além da categoria Espaço, os usuários fizeram uso de muitos assuntos da categoria Tempo para formular as consultas. Em oito das dez consultas, pelo menos um usuário fez uso de um termo da categoria Tempo, o que sugere que a classificação temporal de documentos corporativos seja essencial. Os usuários fizeram uso do tempo para filtrar a informação que sabem ser abundante ou recorrente, para atingir o mais antigo ou o mais recente. Na *Web*, estudos como o de Campos *et al.* (2011) sugerem que menos de 1,5% das consultas façam uso de termos explícitos da categoria Tempo.

Porém, a classificação espacial não se bastaria de um simples atributo ‘local’. Acostumados à busca em texto completo, os usuários fizeram um uso de localidade que pode ser melhor explorada através de uma classificação facetada. Na busca pelo documento 10, os usuários 2, 3 e 4 usaram localidades diferentes. Nesse caso, o evento aconteceu no endereço “expominas”, que não está localizado em “timóteo”. Porém, os usuários 3 e 4 tentaram recuperar um documento que descreve a participação de pessoas de “cefet timóteo” no evento “congresso abm”, enquanto o usuário 2 tenta recuperar um documento que descreve o evento “congresso abm” que aconteceu em “expominas”.

Também, a classificação temporal não se basta de um único atributo ‘data de criação’ pelo mesmo motivo. Para buscar pelo documento 7, os usuários 1 e 2 fizeram uso de dois termos diferentes para a mesma categoria Tempo, “2009” e “2010”. O uso de anos diferentes se explica ao observar a ordem em que foram usados na expressão de busca. O primeiro como “pró-técnico 2009”, o segundo como “pró-técnico vestibular 2010”, significando que o “pró-técnico” aconteceu em 2009 para um vestibular que ocorreria em 2010 ou cujo ingresso ocorreria em 2010. Isso remete à importância de compreender como os usuários de informação constroem a mensagem que a consulta carrega, associando corretamente tempo e espaço às entidades certas por meio de uma classificação facetada e de uma infraestrutura de dados espaciais, local ou corporativa (ALVES; DAVIS JR, 2006).

De fato, se a informação corporativa e as buscas e seus usuários reúnem tantas particularidades que não são compartilhadas com outros contextos de busca, como da própria *Web*, são necessários métodos e tecnologias específicos para tornar o processo de recuperação mais eficaz e eficiente. Na próxima seção são apontadas conclusões, limitações deste trabalho e algumas direções de trabalho futuro.

6 CONCLUSÕES E TRABALHOS FUTUROS

Este trabalho avaliou uma coleção de documentos corporativos e um conjunto de consultas de seus usuários para identificar características comuns a ambos. Oito categorias fundamentais foram propostas e verificadas como características comuns a documentos e consultas, e a distribuição de assuntos entre as categorias apresentou alta correlação positiva. Essas categorias têm o potencial de tornarem-se facetas de uma classificação facetada para a informação corporativa. Isso sugere um caminho para a organização e representação do conhecimento e da informação corporativa que não foi identificado na literatura.

As oito categorias constituem possivelmente um modelo de longo prazo que requer revisões menos frequentes e suporta o desenvolvimento incremental de sistemas de recuperação de informação mais eficientes. O modelo facetado beneficia-se da análise facetada para o reconhecimento de características úteis da informação corporativa, em uma empresa ou em todo o domínio corporativo, algo essencial para o desenvolvimento de produtos de informação generalistas ou especializados.

Os resultados desta pesquisa podem ser comparados facilmente com pesquisas realizadas com dados de outras empresas, sem que os dados sejam muito expostos. O método facetado mostrou-se adequado para esse objetivo ao comparar assuntos e categorias de assuntos de mais alto nível, ao invés de termos e classificações que exponham potencial informação. Adicionalmente, este trabalho constitui um exemplo de classificação facetada mais distante daqueles encontrados na literatura, o que, para Wild *et al.* (2009) e La Barre (2010), é necessário para validar a própria análise facetada fora do seu contexto original.

Até onde sabemos, este é o primeiro trabalho que tenta estabelecer facetas comuns da informação corporativa no conteúdo de documentos e nas consultas de usuários. Portanto, as conclusões são preliminares e requerem amostras e resultados adicionais para que possam ser generalizadas. São igualmente importantes amostras de novos grupos de usuários, cenários adicionais de trabalho e tomadas de decisão, assim como também exemplos de outras empresas. Em todos os casos de novas amostras, as diferenças em tarefas, terminologia, conhecimento prévio sobre a coleção, formação acadêmica, atividade econômica da empresa e idioma devem ser especialmente investigadas, pois são as variáveis com maior potencial de influência sobre as categorias que foram descartadas neste trabalho.

Os efeitos deste modelo facetado em métodos automáticos é outra direção importante de trabalho futuro. A literatura apresenta algoritmos especiais para facetas especiais, tais como algoritmos sociais, espaciais e temporais. Porém, a combinação desses algoritmos em um único e integrado arcabouço é abordada apenas superficialmente devido a concorrência

comercial e ao sigilo industrial, especialmente na área de Ciência da Computação. Portanto, processos automáticos sobre informação corporativa carecem de uma maior base teórica e prática, sendo bons candidatos a objeto de estudo pela Ciência da Informação.

Ainda, coleções de referência são frequentemente adotadas para realizar esses estudos. Apesar de pertinentes para muitos trabalhos que dependam de experimentação, as coleções apresentam sérias limitações à variedade de situações que devem ser estudadas. Este trabalho abre uma última direção de trabalho futuro ao colocar à disposição sua coleção de documentos em língua portuguesa. A coleção pode ser obtida através da Internet e usada sem custos exclusivamente em pesquisa científica.

Espera-se que este trabalho sirva como um guia metodológico para o propósito de uma análise de domínio corporativo. Ao mesmo tempo, caso as categorias propostas se provem perspectivas utilitárias para um número maior de usuários em situações mais diversas, acredita-se também que seus resultados derivados deem suporte ao aperfeiçoamento de métodos automáticos de indexação, classificação e *ranking*.

REFERÊNCIAS

- ABEL, F. *et al.* Leveraging the semantics of tweets for adaptive faceted search on twitter. In: AROYO, L. *et al.* (Ed.). 10th International Conference on the Semantic Web. 10., 2011, **Proceedings...** Heidelberg: Springer, 2011. p.1-17.
- ALBRECHTSEN, H. Subject analysis and indexing: from automated indexing to domain analysis. **Indexer**, v.18, p.219–219, 1993.
- ALVES, L. L.; DAVIS JR, C. A. Interoperability through Web Services: Evaluating OGC Standards in Client Development for Spatial Data Infrastructures. In: Brazilian Symposium on GeoInformatics, 8., 2006. **Proceedings...** São José dos Campos: INPE, 2006.
- ALWIS, S. M. G. de; HIGGINS, S. E. Information as a tool for management decision making: a case study of Singapore. **Information Research**, v.7, n.1, paper 114, 2001.
- AUER, S. *et al.* Linked geodata: Adding a spatial dimension to the web of data. In: International Semantic Web Conference, 8., 2009. **Proceedings...** Heidelberg: Springer, 2009. p.731–746.
- BAILEY, P. *et al.* The CSIRO enterprise search test collection. **SIGIR Forum**, v.41, n.2, p.42–45, 2007.
- BORGES, K. A. V. *et al.* Discovering geographic locations in web pages using urban addresses. In: 4th ACM workshop on Geographical information retrieval, 4., 2007. **Proceedings...** New York: ACM, 2007. p. 31–36.
- BROUGHTON, V. The need for a faceted classification as the basis of all methods of information retrieval. **Aslib Proceedings**, v. 58, n.1/2, p.49–72, 2006.

- BUKOWSKA, E. *et al.* Ontology-based retrieval of experts: The issue of efficiency and scalability within the extraspec system. In: QUIRCHMAYR, G. *et al.* (Ed.). **Multidisciplinary Research and Practice for Information Systems**. Heidelberg: Springer, 2012. p.272–286.
- CAMPOS, R. *et al.* What is the temporal value of web snippets? In: 1st International Temporal Web Analytics Workshop, 1., 2011. **Proceedings...** Hyderabad, India: CEUR, 2011.
- CARDOSO, N.; SANTOS, D. To separate or not to separate? Reflections about current GIR practice. In: Workshop on Novel Methodologies for Evaluation in Information Retrieval, 1., 2008, Glasgow. **Proceedings...** Glasgow, NMEI, 2008.
- CHOO, C. W. *et al.* Information culture and information use: an exploratory study of three organizations. **Journal of the American Society for Information Science and Technology**, v.59, n.5, p.792–804, 2008.
- DAVIS Jr. *et al.* Inferring the Location of Twitter Messages Based on User Relationships, **Transactions in Geographic Information Systems**, v.15, p. 735-751, 2011.
- DAVIS Jr., C. A.; ALVES, L. L. Web Services, Geospatial. In: SHEKHAR, S.; XIONG, H. **Encyclopedia of GIS**. Berlin: Springer, 2007.
- DOLBY, J. *et al.* Extracting enterprise vocabularies using linked open data. In: International Semantic Web Conference, 8., 2009. **Proceedings...** Heidelberg: Springer, 2009. p.779–794.
- GARDIN, J.-C. Document analysis and linguistic theory. **Journal of Documentation**, v.29, n.2, p.137–168, 1973.
- GARFIELD, E. A tribute to SR Ranganathan, the father of indian library science. **Current Contents**, v.6, p.37-44, 1984.
- GIESS, M. *et al.* The generation of faceted classification schemes for use in the organisation of engineering design documents. **International Journal of Information Management**, v.28, n.5, p.379–390, 2008.
- GIRGENSOHN, A. *et al.* Docubrowse: faceted searching, browsing, and recommendations in an enterprise context. In: 15th international conference on Intelligent user interfaces, 15., 2010. **Proceedings...** New York: ACM, 2010. p. 189–198.
- GOPINATH, M. Ranganathan's theory of facet analysis and knowledge representation. **DESIDOC Journal of Library & Information Technology**, v. 12, n. 5, p. 16-20, 1992.
- GUY, I. *et al.* Best faces forward: a large-scale study of people search in the enterprise. In: Conference on Human Factors in Computing Systems. 30., 2012. **Proceedings...** New York: ACM, 2012. p.1775–1784.
- HALEVY, A. Y. *et al.* Enterprise information integration: successes, challenges and controversies. In: International Conference on Management of Data, 31., 2005, Baltimore. **Proceedings...** New York, ACM, 2005. p.778–787.
- HJØRLAND, B. Domain analysis in information science: eleven approaches—traditional as well as innovative. **Journal of documentation**, v. 58, n. 4, p. 422–462, 2002.

- HJØRLAND, B. Facet analysis: the logical approach to knowledge organization, **Information Processing & Management**, v.49, n.2, p. 545-557, 2012.
- HONG, M. Potential usage of faceted classification in internet information retrieval. **Interdisciplinary information sciences**, Tohoku University, v. 12, n. 1, p. 43–51, 2006.
- JONES, R.; HASSAN, A.; DIAZ, F. Geographic features in web search retrieval. In: JONES, C.; PURVES, R. (Ed.). 2nd International Workshop on Geographic Information Retrieval, 2., 2008. **Proceedings...** New York: ACM, 2008. p.57–58.
- LA BARRE, K. Facet analysis. **Annual Review of Information Science and Technology**, v. 44, n. 1, p. 243–284, 2010.
- LIU, X. *et al.* Entity centric query expansion for enterprise search. In: 21st ACM International Conference on Information and Knowledge Management, 21., 2012. **Proceedings...** New York: ACM, 2012. p.1955–1959.
- LIU, X. *et al.* Finding relevant information of certain types from enterprise data. In: 20th ACM international conference on Information and knowledge management, 20., 2011. **Proceedings...** New York, ACM, 2011. p. 47-56.
- LYKKE-NIELSEN, M. Domain analysis, an important part of thesaurus construction. **Advances in Classification Research Online**, v.11, n.1, p.9–50, 2011.
- MARCELLA, R.; ILLINGWORTH, L. The impact of information behaviour on small business failure. **Information Research**, v.17, n.3, paper 525, 2012.
- NUNES, M. B. *et al.* Knowledge management issues in knowledge-intensive SMEs. **Journal of Documentation**, v. 62, n. 1, p. 101–119, 2006.
- O’FARRILL, R. T. Information literacy and knowledge management at work: Conceptions of effective information use at nhs24. **Journal of Documentation**, v.66, n.5, p.706–733, 2010.
- OREN, E. *et al.* Extending faceted navigation for rdf data. In: International Web Semantic Conference, 5., 2006, Atenas. **Proceedings...** Heidelberg: Springer, 2006. p.559–572.
- OVERELL, S. **Geographic Information Retrieval: Classification, Disambiguation and Modelling**. 2009. 181 p. Tese (Doutorado em Computação) - Imperial College, London, 2009.
- RANGANATHAN, S. R. **Prolegomena to Library Classification**. New York: Asia Publishing House, 1967.
- RULA, A. *et al.* On the diversity and availability of temporal information in linked open data. In: International Web Semantic Conference, 11., 2012, Boston. **Proceedings...** Heidelberg: Springer, 2012. p.492–507.
- SACCO, G. M. Some research results in dynamic taxonomy and faceted search systems. In: 18th International Workshop on Database and Expert Systems Applications, 18., 2007. **Proceedings...** New York, 2006. p.201-206.
- SOLOMON, P. Bringing people, technology, and systems together through classification research: Designing, for change, learning, and maintenance. **Advances in Classification Research Online**, v.13, n.1, p.23–28, 2002.

SPINK, A. *et al.* Searching the web: The public and their queries. **Journal of the American society for information science and technology**, v.52, n.3, p.226-234, 2001.

SPITERI, L. A simplified model for facet analysis. **Canadian journal of information and library science**, v.23, p.1-30, 1998.

VAKKARI, P.; JÄRVELIN, K. Explanation in information seeking and retrieval. In: SPINK, Amanda; COLE, Charles (Ed.). **New directions in cognitive information retrieval**. Heidelberg: Springer, 2005. p.113–138.

VICKERY, B. Faceted classification for the web. **Axiomathes**, Springer, v. 18, n. 2, p. 145–160, 2008.

WANG, Z. *et al.* Using classification schemes and thesauri to build an organizational taxonomy for organizing content and aiding navigation, **Journal of Documentation**, v.64, n.6, p. 842-876, 2008.

WILD, P. J. *et al.* Describing engineering documents with faceted approaches: observations and reflections. **Journal of Documentation**, v.65, n.3, p.420–445, 2009.

WU, M. *et al.* Cost and benefit estimation of experts' mediation in an enterprise search. **Journal of the Association for Information Science and Technology**, v.65, n.1, p.146-163, 2014.

ÍCONE, ÍNDICE E SÍMBOLO, FUNDAMENTOS PARA LER E ORGANIZAR A INFORMAÇÃO EM IMAGENS

ICON, INDEX AND SYMBOL, FOUNDATIONS TO READ AND ORGANIZE INFORMATION IN IMAGES

Paulo Roberto Gomes Pato

Resumo: Este artigo apresenta metodologia para a leitura e indexação de imagens. Contrapomos a proposta ao paradigma vigente na Ciência da Informação que enfatiza o processo descritivo e nominativo dos ícones figurados nas imagens. A semiótica de Peirce fundamenta o trabalho. Evidenciar ícones, índices e símbolos é o primeiro passo para extrair a informação presente nas imagens. Compreender as relações entre esses signos semióticos para então extrair algum significado é o caminho lógico de qualquer análise. Portanto, para organizar a informação imagética com eficácia é imprescindível considerar a presença das três principais categorias semióticas de signo, o ícone, o índice e o símbolo, pois os significados emergem apenas em função da semiose, a ação e relação entre esses signos. Para indexar imagens, propomos quatro diferentes entradas de termos. Uma para cada categoria de signos e a quarta para o assunto da imagem. O procedimento mostra que as três categorias de signos formam os predicados do conceito da imagem, que é representado pelo assunto.

Palavras-chave: Imagem. Semiótica. Organização da informação. Indexação. Conceito

Abstract: This article presents a methodology for reading and indexing of images. We oppose the proposal to the prevailing paradigm in information science that emphasizes the process of verbal and naming icons figured the images. The Peircean semiotics based work. Show icons, indexes and symbols is the first step to extract the information present in the images. Understanding the relationships between these semiotic signs to extract some meaning then is the logical analysis of any way. Therefore, to organize information effectively imagery is essential to consider the presence of the three main categories of semiotic sign, icon, index and symbol, because the meanings only emerge as a function of semiosis, the action and the relationship between these signs. To indexing images, we propose four different inputs terms. One for each category of signs and the fourth to the subject of the image. The procedure shows that the three categories of signs form the predicates of the concept image, which is represented by the subject.

Keywords: Image. Semiotics. Organization of information. Indexing. Concept.

1 INTRODUÇÃO

A prática na Ciência da Informação (CI) se manifesta principalmente na organização da informação materializada nos mais diversos suportes: livros, documentos escritos, filmes, fotografias. As características físicas dos documentos imagéticos guardam um primeiro nível de informação, tais como tipo, cor, dimensões. A indexação por esses atributos intrínsecos é baseada em conteúdo. No segundo nível, há um processo de “tradução” da substância intelectual dos suportes e da representação das mensagens por meio de signos. Essa indexação é produzida por humanos e baseada em **conceitos**, e é o foco deste artigo.

Tradicionalmente, a indexação de imagens por conceito enfatiza o processo descritivo. O objetivo principal é o de descrever e nominar ícones semióticos, ou seja, os objetos

figurados em uma imagem. Porém, a semiose, a ação entre os signos que possibilita a constituição de significados, é um processo que envolve necessariamente as três principais categorias semióticas de signo: o ícone, o índice e o símbolo. Entendemos que esses signos, quando identificados em uma determinada imagem e nominados segundo sua categoria, constituem os predicados necessários para a determinação de um conceito – ou assunto – que represente essa imagem. Procuramos mostrar neste artigo que os procedimentos predominantes empregados na indexação de imagens são limitados e insuficientes para evidenciar a informação presente nas imagens.

2 PRINCÍPIOS DOMINANTES NA LEITURA E INDEXAÇÃO DE IMAGENS

Muitos autores desenharam caminhos para a análise e interpretação de imagens estáticas ou em movimento, tais como Baxandall (1991), Gombrich (2007), Wollen (1984), Lindekens (1976), Joly (2012). Porém, segundo Markkula e Sormunen (2000, p. 4), “o trabalho do historiador da arte Erwin Panofsky teve um impacto importante sobre o desenvolvimento teórico na indexação de imagens”. Sua abordagem, embora direcionada às pinturas do Renascimento, está direta ou indiretamente presente nas obras de muitos pesquisadores da área de organização da informação em imagens, como Rasmussen (1997), Kattnig (2002), Shatford Lane (1994; 2002) e Shatford (1986). As duas últimas são a mesma autora, e provavelmente seja a mais conhecida e referenciada na área de leitura e indexação de imagens.

No Brasil, constituiu a base do trabalho de Smit (1997; 1999), que influenciou as pesquisas de Manini (2002), Amaral (2009), Oliveira (2013) e outros. Portanto, parece prudente analisar em linhas gerais a perspectiva de Panofsky (2011), confrontá-la com os postulados de Shatford Lane (1994; 2002), Shatford (1986) e contrapor à tese que defendemos, tendo como pano de fundo a Teoria do Conceito de Dahlberg (1978a; 1978b).

Para Dahlberg (1978b), a evolução dos sistemas de classificação, com base nas classificações facetadas, que, por sua vez, são suportadas pela “Classificação dos Dois Pontos”, de Ranganathan, aliada às pesquisas sobre a elaboração de tesouros, mostrou que as classes, como subdivisões de um todo, não deveriam ser os elementos básicos para a construção de sistemas de classificação. Como decorrência, Dahlberg (1978b, p. 10) afirma que os conceitos, em função de “sua fecunda capacidade de expressão e de combinação” deveriam assumir a posição antes ocupada pelas classes como elementos fundantes dos sistemas de classificação. Prossegue a autora afirmando que um sistema de classificação – enquanto sistema de ordenação – deve ser entendido no sentido de se obter alguma

determinada ordem, quer seja de elementos materiais – documentos, por exemplo – ou de conceitos, que é o que nos interessa neste artigo. Portanto, cabe esclarecer como a autora define conceito e como isso se aplica à nossa proposta de leitura e indexação de imagens, além de servir de base para a análise do que propõe Shatford (1986).

Dahlberg (1978a) explica que muitas coisas são chamadas "definição", e que alguns a entendem como a explanação do sentido de uma palavra. Outros como a simples descrição de um objeto, ou mesmo como processos contidos nos sistemas axiomáticos da matemática e da lógica. Entendemos que o segundo sentido de “definição” elencado pela autora, a descrição de um objeto, é o utilizado atualmente na indexação de imagens, o que questionamos neste artigo. Ou seja, a pura e simples descrição pode “definir” uma determinada imagem fotográfica.

Dahlberg (1978a, p.106) prossegue afirmando que uma “definição equivale a estabelecer uma ‘equação de sentido’”. Em certos termos, a definição é uma limitação, ou seja, a colocação de limites em um conceito ou ideia. Portanto, “definição [é] a delimitação ou fixação do conteúdo de um conceito (conteúdo do conceito = intensão, ou conjunto de características ou atributos).” (DAHLBERG, 1978a, p.106, comentário da autora).

Para se aproximar do objeto de suas especulações – a definição de conceito –, Dahlberg (1978b) diz que desde a primeira infância somos habituados a pensar por meio de conceitos. Enfatiza que isso é conseguido ao se associar o conteúdo dos conceitos a determinados sons ou sinais, e que é natural transportar essas associações para os conceitos científicos, e que assim “confirmamos caráter absoluto a um conhecimento que de nenhum modo o possui.” (DAHLBERG, 1978b, p. 10).

Fundamentada nas formulações do lógico norte-americano Moravcsik (1977)³⁶ sobre a escala de formação dos conceitos, Dahlberg (1978b) afirma que é apenas em função da reflexão e da descrição e do âmbito de aplicação do conceito que surgem as condições para o debate de caráter científico sobre a validade do conteúdo dos mesmos. Conclui afirmando que podemos propor como científicos “somente os conceitos plenamente descritíveis ou definíveis.” (DAHLBERG, 1978b, p. 11).

É importante destacar que, ao abordar a questão dos objetos como elementos do conceito, Dahlberg (1978b) recorre à descrição do processo interno de formação de conceitos

³⁶ MORAVCSIK, J. M. On understanding. In: International Workshop on the Cognitive Viewpoint. University of Ghent, p. 73-82, 24-26 mar., 1977.

de Engelkamp (1976)³⁷, que é baseada nos **objetos da percepção visual**. Segundo esse autor, nossa percepção consiste em uma “representação icônica direta (por meio de imagens ou figuras) do mundo óptico que nos circunda. Esta forma de representação é considerada mais elementar que a constituída pela representação simbólica.” (ENGELKAMP, 1976, *apud* DAHLBERG, 1978b, p. 11, comentário do autor). Ou seja, a representação do mundo passa primeiramente pelo ícone, o figurado, depois pelos índices e posteriormente pelos símbolos, pelas palavras e signos convencionados, como enfatizamos neste artigo e que é o percurso necessário para a indexação de imagens. O autor afirma que a representação icônica é caracterizada por ser concreta e plástica e distribuída em unidades e, nesse sentido, os símbolos (palavras) têm por base os dados da percepção (os ícones). Desses dados são então abstraídos os aspectos possíveis, sejam quais forem – os quais, pela nossa perspectiva de análise de imagens, são os índices semióticos –, **que são então os predicados dos objetos da percepção**. Destaca ainda que o que armazenamos como átomos da estrutura de nosso conhecimento são os predicados, de modo que as unidades elaboradas em nossa memória simbólica, logo, semântica – diríamos semiótica –, em forma abstrata **são configurações desses mesmos predicados**. Assim, os dados percebidos determinam quais predicados devem ser reunidos nas unidades semânticas abstratas. Engelkamp (1976, *apud* DAHLBERG, 1978b, p. 12) conclui afirmando que “os conceitos são feixes de predicados que permitem reunir os dados ou a realidade em classes”.

Dahlberg (1978b, p. 12) amplia a perspectiva de Engelkamp (1976), assentada sobre a experiência visual, e afirma ser possível estendê-la “para outros campos do conhecimento e para outros objetos, sejam eles elementares, como os nossos sentimentos, sejam abstratos, como os produtos do nosso pensamento e das nossas inferências”. A autora ressalva que os predicados devem estar orientados pelo postulado da verdade para poderem ser úteis em uma comunicação intersubjetiva, ou seja, devem ser verificáveis e corresponder à realidade. Apregoa que, quando um predicado possui tal caráter, nele reside um elemento cognoscitivo relacionado com o pensado. Nesse sentido, **o predicado é um elemento do conhecimento**. Portanto, “a reunião dos elementos do conhecimento (predicados) por objeto conduz às unidades do saber [...] Podemos considerar tais unidades como conceitos científicos especiais.” (DAHLBERG, 1978b, p. 12, comentário nosso). Para sabermos algo sobre determinado objeto, no entanto, não precisamos de todos os predicados a ele relacionados, mas somente os predicados necessários. Assim, **a formação dos conceitos situa-se na síntese**

³⁷ ENGELKAMP, J. Satz und bedeutung. Stuttgart: W. Kohlhammer, 1976.

dos predicados necessários verdadeiros a respeito de determinado objeto. Concluindo, define conceito como “a unidade de conhecimento que surge pela síntese dos predicados necessários relacionados com determinado objeto e que, por meio de sinais linguísticos, pode ser comunicado.” (DAHLBERG, 1978b, p. 12). Portanto, podemos formular enunciados de conceitos utilizando as linguagens naturais, os signos semióticos.

Parece claro que, segundo Dahlberg (1978a; 1978b), algumas condições são necessárias para que um objeto qualquer possa ser entendido pelo viés do conhecimento. Para formar o conceito de algum objeto devemos, necessariamente, determinar seus predicados essenciais, os quais, para serem compreendidos e partilhados, devem ser comunicados por meio de entidades linguísticas, ou seja, signos. Portanto, a tarefa fundamental em qualquer processo de classificação e organização da informação constitui-se em organizar o objeto a partir de seus elementos formadores, que por sua vez resultam nos predicados essenciais do conceito desse objeto.

Devemos observar que o conceito é abordado principalmente tendo em vista o viés científico, ou seja, a produção de termos (conceitos) com determinada univocidade em certos domínios de conhecimento. Porém, cremos ser pertinente aplicar na classificação de imagens os princípios relativos ao conceito, pois, de certa maneira, foi por meio do processo de percepção visual que a autora fundou seus princípios gerais. Nesse sentido, afirmamos que o **assunto**, ou **conceito** de uma imagem, é determinado pelo conjunto dos elementos figurados nessa imagem e suas respectivas características. Os elementos figurados e suas características – os **predicados** – são representados pelos ícones, índices e símbolos semióticos. Portanto, para estabelecer e caracterizar o assunto – ou conceito – durante a interpretação de imagens, devemos evidenciar o funcionamento característico de cada um dos três signos semióticos básicos e nominá-los de acordo com suas peculiaridades e forma de funcionamento na semiose.

Embora nossa proposta de leitura e indexação de imagens não esteja vinculada diretamente ao trabalho de Panofsky (2011) ou Shatford Layne (1994; 2002), mas sim à Semiótica de Peirce, avaliamos essas perspectivas de análise para encontrar pontos em comum ou mesmo divergentes e conflitantes com a nossa proposta. Portanto, procuramos destacar como a lógica dos signos atua nas concepções desses autores e como isso reflete em nossa proposta e, conseqüentemente, na organização da informação.

Ao abordar a indexação de imagens estáticas, Shatford Layne (1994) destaca quatro categorias gerais e necessárias de atributos – biográficos, exemplificativos, de relacionamento e de assunto. Afirma que os **atributos de assunto** refletem o que as pessoas pensam quando

indexam uma imagem, sendo, sem dúvida, “uma das mais problemáticas e **menos objetivas** das categorias, e **frequentemente a mais importante.**” (SHATFORD LAYNE, 1994, p. 584, tradução e grifo nossos). Portanto, vamos discorrer apenas sobre esse tipo de atributo, que está vinculado ao conceito, conforme vimos em Dahlberg (1978a; 1978b).

Há três aspectos dos atributos de assunto de imagem que devem ser considerados, segundo a autora. O primeiro é que uma imagem pode ser tanto **DE** e **SOBRE**. Ou seja, é a diferença entre o significante e o significado. Por exemplo, a imagem **DE** uma pessoa chorando pode ser **SOBRE** a tristeza. Portanto, quando a imagem é **DE**, o aspecto mais provável é que a figuração seja concreta e objetiva; quando uma imagem é **SOBRE**, há mais propensão de ser abstrata e subjetiva (SHATFORD LAYNE, 1994). O **DE** está vinculado à **percepção e descrição** do que está representado na imagem: uma pessoa chorando. Esse nível é básico, primário, e está vinculado ao reconhecimento das figuras representadas em uma imagem. O **SOBRE** exige perceber e inferir que a pessoa que chora pode significar tristeza naquela situação.

No segundo aspecto, prossegue Shatford Layne (1994), uma imagem é, simultaneamente, **genérica e específica**. O todo passível de descrição em uma imagem pode ser útil tanto para a sua identidade específica como para a genérica. Como exemplo, mostra a autora, uma imagem da ponte de Brooklyn tanto pode ser útil para um pesquisador que busca uma ponte específica – a de Brooklyn –, como outro que procura uma ponte genérica. Esclarece que **vários termos genéricos** podem ser aplicados a uma única pessoa, objeto ou evento **específico**: a ponte de Brooklyn é uma ponte, ou seja, uma ponte genérica, e é também uma ponte do tipo suspensa, um tipo específico de ponte. Idealmente, “o acesso deve ser fornecido **a todas as possíveis identidades genéricas, bem como a uma identidade específica** de pessoa, objeto, ou evento.” (SHATFORD LAYNE, 1994, p. 584, tradução e grifo nossos). Adverte que uma questão a ser considerada na indexação de imagens é a de indexar cada imagem individual com todos os seus possíveis termos genéricos, ou então criar ligações entre os vários níveis de termos genéricos e específicos em um sistema.

No terceiro aspecto, os assuntos de uma imagem podem ser classificados em **quatro facetas**: tempo, espaço, atividades e eventos, e objetos, tanto os animados quanto os inanimados. Shatford Layne (1994) esclarece que uma imagem pode ser especificamente **DE** e genericamente **DE** ou **SOBRE** qualquer uma dessas facetas. Tomando como exemplo o retrato “*Mrs. Siddons as the Tragic Muse*”, pode-se dizer que a imagem é especificamente sobre a senhora *Siddons* – **DE** –, genericamente sobre uma mulher – **DE** –, e **SOBRE** a musa trágica. Classificando os assuntos das imagens para as facetas acima, e considerando que cada

imagem pode ser **DE** ou **SOBRE** cada faceta, não significa que cada imagem será igual ao **DE** ou o **SOBRE** cada uma das facetas, mas apenas que isso são possibilidades. Portanto, segundo esse raciocínio, o assunto da imagem permanece no campo das possibilidades.

Quanto ao **significado** das imagens, Shatford (1986) esclarece que existem **três níveis**. O primeiro é o **descritivo**, que é o processo de descrição genérica dos objetos e ações representadas na figura. É a descrição factual. A imagem é **SOBRE** o que? Nesse sentido, as descrições do clima da imagem são descrições expressionais. **É no nível descritivo que geralmente se atribui os termos**. Está vinculado ao nível **pré-iconográfico** de Panofsky (2011).

O segundo nível é o **analítico**. Pergunta-se qual é o contexto? Esse nível, segundo Shatford (1986), **requer o conhecimento de uma cultura específica**. Embora possamos manter o contexto em mente, frequentemente **não se é capaz de efetuar uma descrição analítica**. Está vinculado ao **nível iconográfico** de Panofsky (2011).

O terceiro nível é o **interpretativo**. Qual é o significado intrínseco da imagem? Quais sentimentos a imagem traz à mente? Shatford (1986) afirma que esse nível **não pode ser indexado com qualquer grau de consistência**. Propõem que, em geral, **não é apropriado fornecer interpretações sobre imagens**. Está vinculado ao **nível iconológico** de Panofsky (2011).

Esclarecidos os pressupostos de Shatford Layne (2002; 1994; 1986), passamos a considerar o que Panofsky (2011) prescreve para analisar obras de arte. Afirma esse autor que “a iconografia é o ramo da história da arte que trata do tema ou mensagem das obras de arte em contraposição à sua forma.” (PANOFSKY, 2011, p. 47). A partir dessa afirmação esclarece a distinção entre forma e significado. Nesse sentido, explica como identificamos automaticamente o gesto de cumprimento de um homem que tira o chapéu. O que vemos inicialmente, numa pequena fração de tempo quase imediata, é uma configuração geral de cores, linhas e volumes, e isso constitui o nosso mundo visual, diz Panofsky (2011). Em termos semióticos, entendemos que esse momento é a **primeiridade**, instante vinculado à fenomenologia e ao ícone peirceano, o signo que comporta em si as qualidades das coisas. Não há processo propriamente interpretativo, mas percebe-se algum fenômeno plasmado em ícones.

Em seguida, prossegue Panofsky (2011), percebemos a configuração como um objeto bem definido – é um cavalheiro – e a alteração de sua postura, o gesto, como um evento específico – tirar o chapéu – que ultrapassa os limites da percepção puramente formal e adentra a **primeira esfera** da mensagem ou significado. O significado assim percebido é de

natureza elementar e facilmente compreensível. É o **significado fatural**, percebido pela simples identificação de certas formas visíveis vinculadas a determinados objetos familiares em função de nossa experiência prática cotidiana (PANOFSKY, 2011). Podemos vincular esse momento de identificação de algo à **secundidade** peirceana, instante no qual nossa atenção interpretativa é despertada. A secundidade é ligada às ciências normativas e ao índice, o signo semiótico que indica algo sobre as coisas e objetos representados por ícones. Em uma imagem, é o signo que empresta aos ícones os meios de identificação e, posteriormente, algum tipo de identidade.

Objetos e eventos assim identificados produzem uma reação no observador em função de seus conhecimentos e de sua habilidade em perceber sentimentos amigáveis ou hostis, por exemplo, pondera Panofsky (2011). As nuances psicológicas fornecem aos gestos um **significado expressivo**. Enquanto o **significado expressivo** é aprendido, o **significado fatural** não é uma simples identificação, mas uma identificação por “empatia”, a constatação sensível da presença de algo. A compreensão do objeto exige certa sensibilidade – significado fatural – e, mais ainda, a prática com objetos e eventos cotidianos – significado expressional. Assim, afirma Panofsky (2011, p. 48, grifo nosso), “tanto o significado expressional como o fatural podem classificar-se juntos: constituem a classe dos significados **primários ou naturais**”.

Panofsky (2011) destaca que a constatação de que o ato de tirar o chapéu é uma saudação pertence a um **campo diferente de interpretação**. A saudação é típica do mundo ocidental e remanescente de um gesto utilizado pela cavalaria medieval. Ou seja, é um gesto cultural, logo, um signo socialmente construído e partilhado em determinadas comunidades. Portanto, interpretar o ato de tirar o chapéu como um gesto de cortesia é reconhecer nele um significado **secundário ou convencional**, diferente do significado **primário ou natural**. cremos que esse momento reflete a **terceiridade** peirceana, vinculada à metafísica – que trata de compreender a realidade dos fenômenos –, e ao **símbolo** semiótico, o signo que designa algo por convenção, por acordo social quanto ao seu significado. Nesse momento, colocamos em ação todos os mecanismos cognitivos, acionando nosso repertório significativo para “ler” o ambiente, como no exemplo.

Entendemos que a **primeira esfera** de significado está refletida nos procedimentos do atual paradigma de organização da informação de imagens. A leitura e a indexação de uma imagem têm início, necessariamente, com a percepção dos objetos e coisas retratadas em seu interior, suportadas pelos ícones, e prosseguem com a identificação, a descrição e a nomeação desses objetos, o que só é possível pela presença dos índices, os signos que **indicam** o que um ícone **pode ser**. No caso em análise, significa perceber que a figura

humana da imagem é **possivelmente** um cavalheiro, mais pelas roupas e postura que propriamente pelo gesto de cumprimento cavalheiresco. No entanto, o gesto também se vincula ao índice, pois indica algo sobre a pessoa. Porém, como o índice nada afirma sobre alguma ocorrência – apenas indica uma **possibilidade** de sentido –, o que define esse algo, ou esse índice, é a **convenção social** que diz ser o gesto uma forma de saudação cavalheiresca, ou seja, **simbólica, convencional**. Assim, na escala de construção de significado, o gesto é primeiramente um indício, sintoma de algo – é **índice**. Em seguida, assume a condição de **símbolo** semiótico, isso em função de uma convenção social que atribui determinado significado àquele gesto específico. Portanto, é a **semiose**, a ação dos signos quem produz o resultado interpretativo, e não um signo isoladamente. O **ícone** mostra sem dizer, o **índice** indica alguma possibilidade interpretativa sobre o ícone, e o **símbolo** afirma categoricamente, **para aquele que já conhece e reconhece o gesto**, que ele é uma **saudaçãoconvencionada nas sociedades ocidentais**. Nesse sentido, nossa proposta de análise e indexação comporta toda a situação descrita por Panofsky (2011), indo desde o simples reconhecimento e descrição dos objetos presentes em uma imagem até à compreensão da condição simbólica de alguns deles, como o gesto de tirar o chapéu, por exemplo.

O autor esclarece que, além do evento natural que ocorre no espaço e no tempo – o ato da saudação – indicar formas e sentimentos e ser meio para uma saudação socialmente convencional, o **conjunto de todos esses elementos** pode revelar a um observador treinado a “personalidade” do sujeito que reliza o ato. Assim, a tal personalidade “cavalheiresca”, passível de ser percebida, analisada e compreendida, é condicionada pela história social do sujeito, pelo fato dele ser um homem do século XX que se distingue pela trajetória como indivíduo que vê e reage ao mundo. Porém, destaca Panofsky (2011, p. 49), “na ação isolada de uma saudação cortês, todos esses fatos não se manifestam claramente, porém sintomaticamente”. O que é descoberto dessa forma pode ser chamado de **significado** ou **conteúdo intrínseco** e é um significado **essencial**. Os outros dois tipos de significados, o primário ou natural – **significado fatural** – e o secundário ou convencional – **significado expressivo** –, são fenomenais. Portanto, o que Panofsky (2011) afirma ser “o conjunto de todos esses elementos” nada mais é que a semiose, e o resultado dela é o significado ou conteúdo intrínseco da obra.

Após delinear a análise de um ato cotidiano e trivial, Panofsky (2011) aplica os resultados ao processo de análise de obras de arte, distinguindo os mesmos três níveis de significado. O **primeiro nível** é denominado **pré-iconográfico**, e é o tema primário ou natural, subdividido em fatural e expressional. É aprendido inicialmente identificando-se

formas puras: determinadas configurações de linhas e cores, as representações de objetos naturais, de seres humanos, animais, plantas, casas. Depois, identificando as relações mútuas como eventos e, finalmente, percebendo certas qualidades expressivas como a natureza dolorosa de uma pose ou um gesto. O mundo das formas puras, fornecedor de significados primários ou naturais, é chamado de mundo dos motivos artísticos. A enumeração desses motivos corresponde ao trabalho de descrição **pré-iconográfica** da arte. O primeiro nível de significado, o pré-iconográfico, vincula-se semioticamente ao **ícone**, que representa algum objeto, existente ou não no mundo visível, e ao **índice**, o signo que indica uma possibilidade descritiva e significativa sobre o ícone.

No **segundo nível**, a **iconografia** é formada pelo assunto **secundário ou convencional**. Portanto, **simbólico** ao nível semiótico, pois tudo o que é símbolo é convenção. Significa, segundo Panofsky (2011), saber que o reconhecimento de um grupo de pessoas figuradas em uma imagem e sentadas juntas ao redor de uma mesa, em uma determinada ordem, configuração e com certas posturas e atitudes representa a Última Ceia. Figurações desse tipo são chamadas de alegorias, imagens ou histórias, e sua identificação é atribuição da **iconografia**, afirma Panofsky (2011, p. 53). Para o autor, a “iconografia é, portanto, a descrição e classificação das imagens [...] e fornece as bases necessárias para quaisquer interpretações ulteriores. Entretanto, ela não tenta elaborar a interpretação sozinha”. Completa afirmando que a iconografia coleta e classifica a evidência, mas não está apta a investigar a gênese e o significado das evidências, na medida em que é apenas uma das partes que compõem o processo interpretativo.

A **iconologia** é o **terceiro nível** de análise, do qual emerge o significado intrínseco ou conteúdo da obra de arte e é aprendido pela identificação de princípios subjacentes que revelam, por exemplo, uma classe social ou alguma crença religiosa, os quais são qualificados por certa “personalidade” e plasmados em uma obra de arte. Voltando ao exemplo da Santa Ceia de Da Vinci, Panofsky (2011, p. 52) afirma que interpretar tal obra como um grupo de treze pessoas em torno de uma mesa de jantar e que esse grupo de pessoas representa a Última Ceia é considerar as características composicionais e iconográficas como inerentes à obra. Porém, ao compreender tal obra como um documento da personalidade de Leonardo ou da civilização do Renascimento italiano, significa considerar a obra de arte “como um sintoma de algo mais que se expressa numa variedade incontável de outros sintomas, e interpretamos suas características composicionais e iconográficas como evidência mais particularizada desse ‘algo mais’.” Panofsky (2011, p. 52). Assim, a descoberta e interpretação de tais valores “simbólicos” é o **objeto da iconologia**, em oposição à iconografia. Completando, o autor

afirma que a iconologia é uma iconografia que se torna interpretativa. A **iconologia** é, portanto, “um método de interpretação que advém da síntese mais que da análise.” (PANOFSKY, 2011, p. 54).

Podemos depreender do exposto que evidenciar o que há de simbólico em uma imagem não é um ato arbitrário. O símbolo, no sentido semiótico, é o único signo que possui alguma estabilidade informativa e “definitiva”, na medida em que ele é determinado por convenção social, pela história e cultura de uma sociedade. A proposta de Panofsky (2011) procura resgatar valores de um período passado ao qual não temos acesso direto, imediato. Assim, para entender os valores simbólicos de uma obra, o investigador deve se apoiar em uma série de outros documentos que não a obra em si. Deve ter amplo conhecimento sobre o Renascimento, o que é para poucos. Assim procedendo, Panofsky (2011) busca estabelecer um sentido “definitivo” e unívoco para uma determinada obra de arte.

No entanto, nossa proposta de leitura e indexação procura evidenciar o que há de simbólico em uma imagem apenas até o nível iconográfico. **Não pretendemos estabelecer um valor interpretativo simbólico e unívoco para o que uma imagem significa como um todo**, muito menos entender, pela análise da imagem, aspectos psicológicos ou existenciais do autor, ou o espírito de uma época ou nação como sugere o nível iconológico proposto por Panofsky (2011).

A clássica imagem difundida mundialmente da conquista de Iwo Jima, na Guerra do Pacífico, desdobramento da II Grande Guerra, é um exemplo do que queremos mostrar (FIG. 1). Se bem entendemos o que propôs Panofsky (2011), essa fotografia tem **dois momentos simbólicos**. O **primeiro** está relacionado à configuração geral **do ato representado no interior da fotografia**: o hasteamento ou fixação de uma bandeira. O ato é recorrente em batalhas e pode ser encontrado em inúmeras outras imagens e filmes de guerra ou sobre guerra. Ao tomar a posição inimiga, o vencedor finca sua bandeira no mais alto ponto do território conquistado. Torna-se, portanto, emblema da tomada de parte do campo inimigo. O simples ato é entendido simbolicamente como triunfo em um embate porque foi construído socialmente, tal qual o movimento de cumprimentar alguém pelo gesto de tirar o chapéu, como mostra Panofsky (2011). Esses atos estão vinculados ao **nível iconográfico**.

FIGURA 2 - Tomada de Iwo Jima



Fonte 1- Disponível em:
<<http://www.iwojima.com/bond/lflage.gif>>. Acesso em: 6 jun.
2014.

Além do valor **simbólico primário e genérico** – pois serve para qualquer batalha –, construído ao longo do tempo por inúmeras situações semelhantes, há um **segundo valor simbólico**, também construído e convencionado em um determinado momento histórico. Porém, nesse caso, a convenção simbólica excede aquela primeira e interna da imagem em si, e torna-se, ela própria – a fotografia –, o emblema de algo muito maior que a simples vitória pontual em determinada batalha, à qual extrapola. Ela **adquire valor simbólico como totalidade histórica de uma guerra**, representação que é de uma empreitada muito superior ao ato heróico do confronto limitado a determinada batalha. Nesse sentido, sua configuração transcendeu a fotografia em si e foi corporificada em escultura, simulacro “real” do ato concreto capturado pelo fotógrafo (FIG. 2).

FIGURA 3 - Memorial de Iwo Jima, em Arlington, USA



Fonte 2 - Disponível em:

<<http://www.americashistorylady.com/Photo%20Album.html>>. Acesso em: 6 jun. 2014.

Enfim, a fotografia do hasteamento pode ser compreendida como “o triunfo em uma batalha” (nível iconográfico) e como “o triunfo da ‘América’ na guerra do Pacífico” (nível iconológico). No sentido **iconográfico**, podemos utilizar a imagem como **conceito genérico**(simbólico) de um ato de vitória em batalha, com todos os predicados que isso implica – os soldados, o movimento, a bandeira sendo erguida, o topo de uma colina, etc. Portanto, de forma ampla e genérica, embora seja também conceito específico da batalha de Iwo Jima. No sentido **iconológico**, a imagem é representada pelo **conceito específico**(simbólico) da vitória de uma nação, com os mesmos predicados da imagem no sentido iconográfico.

Embora Shatford (1986) afirme que a segunda etapa de exame é **analítica** e **requer o conhecimento de uma cultura específica**, com o que concordamos, discordamos quando afirma **não ser apropriado fornecer interpretações sobre imagens**, ao se referir ao terceiro nível de análise – o iconológico de Panofsky (2011). Cremos haver alguma confusão conceitual na afirmação se a considerarmos semioticamente. Qualquer análise semiótica é interpretação. Definir o significado simbólico no nível iconográfico, o **segundo nível** e no qual qualquer análise deve se deter, **segundo a autora**, implica também em perceber sentidos simbólicos que a simples representação ou descrição não pode afirmar *per se*. Isso fica muito claro no exemplo acima. Se apenas descrevermos a imagem dos homens levantando uma bandeira, estamos eliminando o sentido simbólico, **pois construído socialmente em uma determinada cultura**, importante para a compreensão daquele ato específico. Devemos interpretar o ato como algo específico, e não como um simples gesto de levantar alguma

bandeira. Portanto, a mera descrição pouco diz sobre a imagem, e ela pode ser colocada em uma mesma categoria junto a outra na qual crianças hasteiam uma bandeira na escola. Em ambas a descrição da situação em si é muito próxima, porém com sentidos distintos, pois seus predicados são diferentes. O aspecto simbólico, também presente no hasteamento na escola, atua como predicado discriminante em ambas as situações.

3 PROPOSTA DE LEITURA E INDEXAÇÃO DE IMAGENS

A seguir vamos exemplificar nossa proposta de leitura e indexação de imagens. Tomamos por base a fotografia “Lock at Bougival”, de 1955, registrada por Henri Cartier-Bresson (FIG. 3).

FIGURA4 - Lock at Bougival, France, 1955 (Henri Cartier-Bresson)



Fonte 3: Disponível em: <http://www.1stdibs.com/art/photography/henri-cartier-bresson-lock-at-bougival-france/id-a_4225/>. Acesso em: 13 fev. 2012.

Todos os elementos estão em harmonia na FIG. 3. Os cães são amigáveis, particularmente aquele que olha diretamente para o homem, pois o outro dirige sua atenção ao fotógrafo. As expressões corporais do cão que mira o homem evidenciam familiaridade possível entre os dois. Essa dedução só é pertinente para aqueles que reconhecem a atitude e o comportamento de animais amistosos. A criança sorri, juntamente com as mulheres. Nota-se o destaque de três elementos pela incidência da luz: o bebê, o torso do homem e um dos

cachorros. Sem considerar os outros elementos da imagem, tão somente essa triangulação – homem, bebê e cão amistoso – já sugere tendência a uma conclusão geral: familiaridade. Familiaridade decorre de família, significando algo que não é estranho, mas sim íntimo. Avançando na leitura, a noção de familiaridade decorrente da recepção amistosa da criança e do cachorro ao homem ganha força pela presença das mulheres. Ambas sorriem e a que parece com mais idade – avó do bebê e sogra ou mãe do homem? – se dirige diretamente ao nenê como se o incitasse a se relacionar com o homem. Logo, a mulher mais jovem deve ser a mãe da criancinha e esposa do homem. Portanto, a esses ícones que emitem sinais (índices) que indicam familiaridade são acrescidos outros que os reforçam, somando dados que melhor ajustam o sentido. Alguns outros elementos se destacam e corroboram a leitura. O barco que está atracado parece ser a casa da família. Deduzimos em função de alguns índices. As cordas amarradas no atracadouro sugerem o barco. A inclinação da cobertura sugere o telhado de uma casa. Mesmo se não soubéssemos que a fotografia foi produzida na França, poderíamos deduzir que era verão, um período de clima mais ameno que o rigoroso inverno europeu. As roupas leves e pés descalços revelam essa possibilidade, e sugerem também que as pessoas estão em ambiente de pouca ou nenhuma formalidade, não são ricas nem miseráveis e estão felizes. O bebê parece saudável, bem cuidado e alegre.

O resultado da análise aponta o **símbolo** convencionalizado “**família**” como uma possível conclusão lógica para a semiose ocorrida em função da ação dos signos identificados na imagem. O símbolo (conceito) “família” evidencia o fato de que apenas ele pode afirmar algo e ser um argumento, **uma conclusão lógica para a interpretação**. Os elementos icônicos e indiciais somente mostram e sinalizam ao intérprete possíveis percursos a serem trilhados, apresentando-se como **premissas** básicas para a análise. Ajustes devem ocorrer, possibilidades são descartadas e consideradas na medida em que a análise transcorre. A presença de certos signos reforça ou rechaça alguma tendência interpretativa. Portanto, ler imagens é considerar a relação global entre os signos, ajustando os sentidos possíveis em função dessa relação e de nosso conhecimento.

O **símbolo** é o único signo que pode ser um argumento. O **argumento** é um signo que no interpretante – na mente do intérprete – é signo de lei; portanto, convencionalizado. O **argumento** (família) é um signo que é compreendido como representando um **objeto** em seu caráter de signo (a imagem fotográfica em seu todo significativo). Portanto, o termo simbólico convencionalizado **FAMÍLIA** representa a imagem fotográfica em sua totalidade, sendo assim seu **CONCEITO**. O argumento, por ser símbolo, deve envolver um signo dicente. O **signo dicente** é um signo que no interpretante – mente do intérprete – é um signo de existência real.

Ele é a **premissa e o predicado** do símbolo, uma vez que um argumento só pode destacar a “lei” sublinhando-a em um caso particular (PEIRCE, 2008).

A totalidade do significado da imagem pode ser resumida na seguinte ponderação: um símbolo é um signo que está naturalmente apto a declarar que um grupo de objetos, indicado por um conjunto de índices que pode estar em certos aspectos ligados a ele, é representado por um ícone associado ao símbolo (CP 2.295). Na análise em questão, e invertendo a afirmação acima, devemos percorrer o texto imagético destacando primeiramente os “objetos” na imagem, que são representados pelos ícones – mulheres, cães, homem, bebê, etc. Esses ícones são signos dicentes, são “reais” para o analista. Como os índices devem necessariamente estar atrelados a ícones para indicar alguma possibilidade significativa, os dicentes – ícones “reais” – carregam em si índices que apontam e sugerem possíveis caminhos para a análise. É bom frisar que os ícones e índices são também, em última análise, símbolos, signos representados por palavras (signos convencionados) que possuem estabilidade significativa a partir da qual podem admitir outros significados. A trama entre os ícones nominados por um símbolo (palavra) – cães, homem, etc. – em coesão com os índices, nominados por um símbolo (palavra) – pescador, marido, etc. –, espalhados por todos os ícones da imagem levam a mente do intérprete a uma conclusão lógica possível, a outros símbolos específicos daquela figuração – pai, marido, mãe, etc. O resultado final é resumido pelo símbolo genérico FAMÍLIA, que é o argumento.

A FIG. 4 sintetiza em linhas gerais os signos que podem ser **observados, detectados, deduzidos e afirmados** na FIG. 3.

FIGURA 5 - Signos da FIGURA 3 e as quatro entradas de palavras-chave

ÍCONES mostram	+	ÍNDICES indicam	+	SÍMBOLOS afirmam	ASSUNTO CONCEITO
		Trabalhador braçal			
Homem		Estivador			
Roupa (macacão e boné)		Pescador			
Torso seminu		Pai do bebê		Pai	
Mãos na cintura		Marido da jovem mãe		Marido	
Visto por trás		Parente			
Olhando mulheres		Filho da senhora			FAMÍLIA
Olhando bebê		Genro da senhora			
		Branco			
		Jovem			
Senhora sorrindo		Mãe jovem			
Mulher descalça		Esposa do homem		Mãe	
sorrindo		Jovem		Avó	
Vestido		Avó			

ÍCONES mostram	+	ÍNDICES indicam	+	SÍMBOLOS afirmam	ASSUNTO CONCEITO
Lenço no cabelo Partes do corpo desnudas		Meia idade Trabalhadoras do lar Parentes Mãe e filha/sogra e nora Brancas			
Bebê Corpo nu		Filho (a) Criança Bebê Neto Branco		Filho (a)	
Cão olhando homem Cão olhando fotógrafo Coleira Orelhas e rabos abaixados		Mansos Amistosos Familiars		Cães domésticos	
Barco de madeira Pier Portas Janelas Forma Cobertura (telhado)		Barco casa Atracado		Residência Lar	
TERMOS GERAIS predicados		TERMOS ESPECÍFICOS predicados		TERMOS ESPECÍFICOS Predicados essenciais da FIGURA 3	TERMOS GERAIS conceito
Homem Mulher jovem Mulher meia idade Bebê Dois cães Barco		Estivador Pescador Marido Jovem esposa Mãe bebê Avó Sogra Nora Nenê Barco casa Cães mansos Cães amistosos		Pai Marido Mãe Avó Filho (a) Cães domésticos Residência Lar	FAMÍLIA

Fonte: Do autor

É interessante destacar que encontramos na *web* uma análise em francês da mesma imagem. O analista avançou um ponto a mais do que revela nossa leitura. Nessa análise, a figura e a postura do bebê na fotografia remetem às clássicas imagens pictóricas do Menino Jesus desde o Renascimento. O analista conclui então com o termo “sagrada família”. Certamente os leitores franceses estão mais próximos e imersos nos símbolos da pintura

renascentista que os leitores brasileiros. Essa conclusão reflete o nível **iconológico** de Panofsky (2011), no qual a fotografia **como um todo** apresenta sentido **simbólico**.

Na análise que empreendemos, os **ícones** funcionam como **termos genéricos** e **predicados** não essenciais do **conceito família**, e se apenas eles fossem indexados teríamos algo como um homem, duas mulheres, bebê, dois cães e barco. Ou seja, não é possível inferir que formam uma **família** em função da generalidade. Tomando a figura do homem como exemplo, temos a seguinte situação: ele é um **ícone genérico**, homem. Um homem pode assumir uma série de **configurações específicas**. Nesse caso, os **índices** indicam (possibilidade) que ele pode ser um **trabalhador**, mais especificamente um **estivador** – está em um píer – ou um **pescador**, pois mora em um barco, está em um píer e usa roupas leves e características de algum dos possíveis tipos de pescador. Deve ser **pai**, pois há uma **criança** e uma **jovem mulher** com a criança se dirigindo a ele. Portanto, os **índices** sugerem **termos um pouco mais específicos** recortados do universo dos possíveis termos naquele contexto e diante da ação de todos os outros signos. Os **termos específicos** representados pelos **símbolos** (na terceira coluna) são os predicados essenciais para o resultado geral da análise, da ação entre os demais signos. O termo **FAMÍLIA**, como conceito final resultante da análise, é novamente um **genérico**, como homem. Assim, para diferenciar essa **família** dos milhões de outras imagens possíveis com essa definição, **devemos considerar os termos como um todo**, mas cada qual em sua classificação específica. Quando da busca por **família**, a presença diferenciadora de **ícones** e **índices** indicará resultados mais recortados e diferenciados. Ou seja, eles **formam os predicados do conceito família** da imagem em análise. Os ícones e índices apontam para os vários objetos da cena e suas possibilidades significativas. Porém, os **símbolos** na terceira coluna formam a essência (os predicados essenciais) do conceito genérico e conclusivo **FAMÍLIA**.

FIGURA 6 - Fotografia de família nos EUA, de Walker Evans



Fonte 5 - Disponível em: <<http://2.bp.blogspot.com/-Im3AA6llaJ4/TY-5ObYIH2I/AAAAAAAAAp5k/8xfiFIAbj0M/s1600/trodd.jpg>>. Acesso em: 14 abr. 2013.

A contraposição entre a FIG. 3 e a FIG. 5 evidencia diferenças. Ambas retratam, de modo geral, um grupo familiar com **ícones semelhantes**. Os **símbolos**, os predicados essenciais do conceito **família**, também são semelhantes: pai, mãe, filhos, etc. Porém, **os índices apontam para conclusões diferentes**. A ambiência revela pobreza, sofrimento, desolação e resignação na FIG. 5. Ambos os ambientes são pobres, mas contrastantes. O da FIG. 3 é, está ou parece feliz. O da FIG. 5, é, está ou parece infeliz. É, está ou parece significa que o registro fotográfico apenas **indica** possibilidades significativas. **Pela foto inferimos tão somente em função do que percebemos**. Apesar da situação, as famílias podem ser felizes ao seu modo no dia a dia, mas isso não podemos saber apenas pelas imagens, e não há garantia alguma sobre isso, obviamente, pois mesmo uma afirmação oral categórica pode ser mentirosa.

4 CONCLUSÃO

As análises que empreendemos exemplificam nossa proposta de leitura e indexação de imagens e confirmam o que defendemos neste artigo. A atribuição do assunto – ou conceito – de uma imagem deve ser definida em função dos três níveis de ocorrência dos três diferentes tipos de signos, pois é a presença diferenciada entre eles que garante o significado global e pode possibilitar o refinamento da busca. Se “família” é o signo comum que une essas imagens (conceito, descritor, o assunto, a conclusão final da semiose), os outros signos são os discriminantes e formam os predicados necessários do conceito. Há quatro níveis de entrada

das palavras-chave, como vimos na Tabela 4, os três primeiros vinculados a cada um dos três tipos de signos e a quarta referente ao conceito ou assunto da imagem. A busca em um banco de imagens que prescrevesse apenas “família” certamente recuperaria as duas imagens do nosso exemplo, pois esse é o conceito ou assunto principal em ambas. Porém, o usuário que buscasse por “família pobre, tristeza, miséria”, baseado em uma indexação como propomos, certamente recuperaria apenas a FIG. 5.

Acreditamos que a indexação se torna mais efetiva se estabelecermos os **quatro níveis separados de entrada de termos**, como o proposto neste artigo. Após as entradas, devemos tabular os termos recorrentes em cada nível – no caso de indexação colaborativa – para em seguida organizá-los ontologicamente como conjunto representativo de uma determinada imagem. Esse conjunto abriga o **conceito** e seus **predicados essenciais e não essenciais**. Inseridos os dados em um sistema computacional, o usuário poderá recuperar uma imagem especificando e cruzando os termos que foram coletados separadamente e inseridos em conjunto e vinculados a uma imagem no sistema. Além disso, permite ao usuário resgatar imagens com o mesmo conceito e predicados diferentes. Assim, o **conceito família**, por exemplo, pode ser representado em todas as **acepções sociais possíveis**, cada qual com seus pertinentes **predicados**.

REFERÊNCIAS

AMARAL, Luciana. **A importância do tratamento intelectual das fotografias visando à recuperação da imagem**. 2009. 141f. Dissertação (Mestrado em Ciência da Informação) – Escola de Comunicações e Artes da Universidade de São Paulo, 2009.

BAXANDALL, Michael. **O olhar renascente: pintura e experiência social na Itália da Renascença**. Rio de Janeiro: Paz e Terra, 1991.

DAHLBERG, Ingetraut. Fundamentos teórico-conceituais da classificação. **Revista de Biblioteconomia de Brasília**, v. 6, n. 1, p. 9-21, 1978b. Disponível em: <<http://www.brapci.ufpr.br/documento.php?dd0=0000008680&dd1=079ad>>. Acesso em: 10 jul. 2013.

DAHLBERG, Ingetraut. Teoria do conceito. **Ciência da Informação**, v. 7, n. 2, p. 101-107, 1978a. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/viewFile/1680/1286>>. Acesso em: 10 jul. 2013.

ENGELKAMP, J. **Satzundbedeutung**. Stuttgart: W. Kohlhammer, 1976.

GOMBRICH, E. H. **Arte e ilusão: um estudo da psicologia da representação pictórica**. São Paulo: Martins Fontes, 2007.

JOLY, Martine. **Introdução à análise da imagem**. Campinas: Papyrus, 2012.

KATTNIG, Cecile. **Gestion et diffusion d'un fonds d'image**. Paris: ADBS, Nathan, 2002.

LINDEKENS, René. **Essai de sémiotique visuelle**. Paris: Klincksieck, 1976.

MANINI, Míriam Paula. **Análise documentária de fotografias**: um referencial de leitura de imagens fotográficas para fins documentários. 2002. Tese (Doutorado em Ciência da Informação e Documentação) - Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2002. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/27/27143/tde-23032007-111516/>>. Acesso em: 14 out. 2012.

MARKKULA, Marjo; SORMUNEN, Eero. End-user searching challenges indexing practices in the digital newspaper photo archive. **Information Retrieval**, Department of Information Studies, University of Tampere, Finland, v. 1, p. 259-285, 2000.

MORAVCSIK, J. M. On understanding. In: **International Workshop on the Cognitive Viewpoint**. University of Ghent, p. 73-82, 24-26 mar. 1977.

OLIVEIRA, Ronni. **Fundamentos da gestão da informação em imagens**: para bibliotecários, arquivistas, museólogos e outros profissionais da informação. São Paulo: Projeto Informação Audiovisual, 2013.

PANOFSKY, Erwin. **Significado nas artes visuais**. São Paulo: Perspectiva, 2011.

PEIRCE, Charles S. **Semiótica**. São Paulo: Perspectiva, 2008.

PEIRCE, Charles S. **The Collected Papers of Charles Sanders Peirce**. Disponível em: <http://www.4shared.com/document/oRnzQCug/The_Collected_Papers_of_Charle.html>. Acesso em: 20/01/2012. (volume 1-8, citado CP seguido pelo número do volume e número do parágrafo).

RASMUSSEN, Edie M. Indexing images. **Annual Review of Information Science and Technology (ARIST)**, Medford, NJ, v. 32, p. 169-196, 1997.

SHATFORD LAYNE, Sara. Some issues in the indexing of images. **Journal of the American Society for Information Science**, v. 45, n. 8, p. 583-588, 1994. Disponível em: <<https://www.yumpu.com/en/document/view/19168757/layne-sara-shatford-1994-some-issues-in-the-indexing-of-images>>. Acesso em: 12 dez. 2013.

SHATFORD LAYNE, Sara. Subject access to art images. In: BACA, Murtha. Introduction to art image access: issues, tools, standards, strategies. **Getty Research Institute**, 2002. Disponível em: <<http://d2aohiyo3d3idm.cloudfront.net/publications/virtuallibrary/0892366664.pdf>>. Acesso em: 02 jun. 2014.

SHATFORD, Sara. Analyzing the subject of a picture: a theoretical approach. **Cataloging & Classification Quarterly**, v. 6, n. 3, p. 39-62, 1986.

SMIT, Johanna W. **A análise da imagem**: um primeiro plano. São Paulo: Mimeo, 1999.

SMIT, Johanna W. Propostas para a indexação de informação iconográfica. **Congresso Internacional de Informática. INFO'97**. Cuba, outubro de 1997.

WOLLEN, Peter. **Signos e significação no cinema**. Lisboa: Livros Horizonte, 1984.

PADRÕES LINGUÍSTICOS PARA EXTRAÇÃO AUTOMÁTICA DO OBJETIVO

A LINGUISTICS PATTERNS FOR AUTOMATIC EXTRACTION OF THE OBJECTIVE

Leonardo Cruz Costa
Luciana Reis Malheiros

Resumo: Este trabalho apresenta padrões linguísticos que auxilia a extração automática do objetivo de artigos científicos na área de biomedicina. Para identificar os padrões foram realizadas análises textuais em cada artigo selecionado, formando um conjunto de 72 (setenta e dois) artigos. A análise serviu para selecionar as frases indicativas (*cue frases, gold-frases*) usadas pelos autores para anunciar o objetivo do seu trabalho. Foram identificadas 78 (setenta e oito) frases indicativas formando 6 grupos linguísticos. A partir disso foi construído um software que permitiu realizar três experimentos. No primeiro experimento foram processados 50 artigos e os resultados foram avaliados por um grupo de pesquisadores. Os resultados obtidos possibilitaram aperfeiçoamentos no software. Um novo experimento foi conduzido no qual outros 60 artigos foram processados sendo obtidos 91,36% para a revocação e 80,30% para a precisão. Finalmente, um terceiro experimento foi realizado com autores reais indicando que eles concordavam que o software encontrou, em 100% dos casos, as frases que anunciavam os objetivos nos seus respectivos artigos científicos.

Palavras-chave: Padrão linguístico. Reconhecimento de padrões. Artigo científico.

Abstract: This work has linguistics patterns that helps the automatic extraction of the objective of biomedical scientific articles. To identify the patterns it was used text analysis in each selected article, making 72 (seventy two) articles. The analysis was used to select the indicative sentences (*cue frases, gold-frases*) which were used by the writers to announce the objective (goal) of their work. There were identified 78 (seventy eight) indicative sentences making six languages groups. After that a software was built allowing to make three experiments. On the first one 50 articles were processed and the results were verified for a group of researchers. The obtained results allowed to polish the software. A new experiment was conducted and 60 other articles were processed. They obtained 91,36% of revocation and 80,30 to precision. Finally, a third experiment was made with real writers and it demonstrate that they consent that the software found, in 100% of the cases, the sentences that had the objective in their respective scientific articles.

Keywords: Linguistics patterns. Automatic extraction. Paper.

1 INTRODUÇÃO

Por meio do artigo científico o pesquisador repassa às comunidades científicas informações e conhecimentos gerados, permitindo que os resultados possam ser confirmados e sua validação também possa ser comprovada através da publicação em jornais e revistas científicas, após julgamento de seus pares (*peer review*). Hoje a publicação eletrônica é uma realidade; apoiando-se em tecnologias tornou-se possível a criação de repositórios para publicação (revistas eletrônicas, *Eprints, preprints, self-publishing*, etc.). Contudo, a Recuperação da Informação continua baseada nas tradicionais máquinas de busca e nos

Sistemas de Informação, onde se gasta uma quantidade significativa de tempo pesquisando, “folheando” e lendo documentos, para descobrir como estão relacionados entre si ou quando apresentam informações que se afastam do objetivo pesquisado. Apesar do formato digital, artigos enquanto bases de conhecimento são voltados para leitura e processamento do seu conteúdo por pessoas.

Artigos científicos são cópias digitais das publicações em papel e não incorporam as potencialidades previstas, como por exemplo, na Web Semântica (MARCONDES, 2005), uma extensão da web atual, onde a informação possui um significado bem definido permitindo melhor interação entre os computadores e as pessoas (BERNERS-LEE, 2001).

A web semântica apresenta diversas e diferentes visões (PASSIN, 2004), e entre elas, a de que deve permitir que os dados disponíveis e “ligados” na internet possam ser usados por máquinas não só para propósitos de exibição, mas para automatização, integração e uso por várias aplicações. E ainda, ela deve ter como função, a criação de uma web que possibilite o uso de agentes inteligentes para que se possa recuperar e manipular a informação de maneira automática e precisa utilizando o contexto no qual ela está inserida ao invés unicamente de palavras-chave.

O princípio básico da web semântica é tornar possível expressar e representar a semântica dos dados através de outros esquemas descritivos, objetivando a precisão dos significados, ou seja: um esquema que descreva quais os conceitos existentes em certo domínio e como eles se relacionam. Como um modelo conceitual, que serve como uma representação abstrata dos elementos de informação. Esse esquema é representado numa ontologia. Observa-se por meio dela como os metadados representam explicitamente a semântica das informações. Portanto, as ontologias têm um papel crucial para web semântica, no sentido de que permitem o acesso, a interoperabilidade e a comunicação baseados em conteúdo.

A ontologia é similar a um dicionário, taxonomia ou glossário, porém com estrutura e formalismo que possibilita computadores processar seu conteúdo. Ela consiste de um conjunto de conceitos, axiomas e relações, e representa uma área de conhecimento. Diferente da taxonomia ou glossário, permite estabelecer relações arbitrárias entre conceitos, propriedades lógicas e semânticas de relações tais como: transitividade, simetria e inferência lógica sobre as relações.

Diante disso Marcondes (2005) propõe uma ontologia para representar o conhecimento contido nos artigos científicos produzidos pela área de biomedicina em formato “inteligível” por programas. De acordo com a ontologia, artigos científicos seriam publicados

não só em formato textual, mas também como uma instância dessa ontologia, representando o conhecimento específico contido em cada artigo. Assim, essa ontologia é usada como modelo de representação e tem como componentes elementos da metodologia científica como: objetivo, problema, hipótese, metodologia, resultados e discussão. Esses elementos semânticos devem ser extraídos do texto do artigo para serem registrados como parte de uma instância da ontologia. A partir da extração e do registro dos elementos semânticos será possível realizar inferências de forma a melhorar a recuperação e relacionamentos de diferentes artigos. Parte dessa ontologia foi integrada em um ambiente de publicação (COSTA, 2011) que interage com o autor e com o texto do artigo, extraíndo e representando o conhecimento. O papel do autor nesse ambiente é, não só validar os elementos semânticos extraídos automaticamente como, também resolver questões de ambiguidade.

Este trabalho tem como fim apresentar padrões linguísticos utilizados para a identificação de partes do texto que se relacionam com os elementos da ontologia, mas especificamente com o objetivo do artigo científico.

2 REFERENCIAL TEÓRICO

A Ontologia do Conteúdo de Conhecimento de Artigos (OCCA) (MARCONDES, 2005) é um modelo para representar os elementos semânticos que constituem o conhecimento contido em um artigo científico, baseado em elementos do Método Científico (método hipotético-dedutivo) e na forma como eles aparecem em artigos científicos, são representados por metadados semânticos. A finalidade desse modelo é servir de base para publicações semânticas, uma forma diferente de publicar artigos científicos. A publicação semântica fornece uma maneira dos computadores compreenderem a estrutura e até mesmo o significado das informações publicadas, tornando a pesquisa de informações e a integração de dados mais eficiente em função da utilização dos metadados semânticos, representados como uma instância de OCCA.

Por meio de OCCA é possível representar quatro diferentes tipos de artigos (Teóricos, Experimentais-exploratórios, Experimentais-indutivo e Experimentais-dedutivos), cada um baseia-se em diferentes raciocínios, estratégias de argumentação e pressupostos observados em diferentes tipos de artigo científico (MARCONDES et al., 2009).

A proposta de OCCA se baseia no entendimento de que o conhecimento científico consiste em propor e provar a existência de relações entre fenômenos e relações entre fenômenos e a suas características (BUNGE, 1998; MILLER, 1947), até então desconhecidas. OCCA é constituído de elementos semânticos como a questão de pesquisa, a hipótese e a

conclusão. Além disso, a conclusão e a hipótese podem ser representada como uma **RELAÇÃO**.

Uma **RELAÇÃO**, para OCCA, tem a forma de um antecedente, uma relação propriamente dita e um conseqüente, como uma 3-tupla (antecedente, relação, conseqüente). Por convenção usa-se a palavra **RELAÇÃO** (em maiúscula) ao se referenciar 3-tupla (antecedente, relação, conseqüente) e *relação* (em minúscula) ao se referenciar a relação dentro da 3-tupla.

A relação pode ser um tipo de relação específica como “causa”, “afeta”, “indica”, ou um tipo de relação **tem_característica** que define uma característica do antecedente, enquanto o antecedente e o conseqüente referem-se aos fenômenos estudados pelo autor.

Enquanto a questão de pesquisa se caracteriza como uma pergunta de caráter geral e não como relação, ao contrário das hipóteses, ao propor uma relação entre fenômenos, têm importância decisiva quanto à manifestação do conhecimento novo em ciência.

Um exemplo disso é o seguinte enunciado: The aim of this study was to assess the cytotoxicity and the potential antiviral activity of violacein against the viruses: Herpes Simplex Virus type 1 (HSV-1) strains KOS, 29-R/acyclovir resistant and VR733/ATCC; Poliovirus type 2 (PV-2); Simian rotavirus SA11 strain, Hepatite A virus (HAV) strains HAF203 and HM175 and Adenovirus type 5 (AdV-5), a respiratory strain, a hipótese é de certa forma "operacionalizada" por meio de um objetivo, em que o autor explicita e enfatiza o fenômeno ou fenômenos, sobre os quais o artigo vai tratar, no exemplo: violacein e vários tipos de vírus. Assim, a seguinte **RELAÇÃO** poderia ser descrita: <violacein> <previne> <vários tipos de vírus>, onde previne, é a relação tem o sentido de parar, impedir ou eliminar uma ação ou condição (como uma atividade antiviral).

A importância de um esquema como de OCCA que utilize relações está no fato de que estas permitem que programas façam “inferências” sobre o conhecimento assim representado, como nos seguintes exemplos: *Violacein* (antecedente) previne (relação) que outras infecções (conseqüente?)? Que outras substâncias (antecedente?) previnem (relação) herpes (conseqüente?) além *violacein*?

Outro aspecto inserido no contexto de OCCA é que a **RELAÇÃO** pode, também, ser mapeada para ontologias públicas, estabelecendo uma ligação entre o artigo e o conhecimento público disponível e validado. A falta de mapeamento entre os elementos da **RELAÇÃO** e a ontologia pode indicar evidências de novas descobertas por causa da falta de representação de tais elementos na ontologia (MALHEIROS, 2006).

Portanto, em OCCA, a questão de pesquisa, a hipótese e a conclusão são elementos semânticos que necessitam ser extraídos dos artigos e representados para futuramente serem usados na recuperação da informação. E nesse sentido, os objetivos são justamente a construção das metas necessárias para se transformar as respostas hipotéticas em respostas reais, concretas, ao serem atingidos comprovam ou negam as hipóteses anteriormente formuladas no contexto da pesquisa, por isso usa identificação e extração no texto é fundamental para o modelo OCCA.

2.1 Artigo Científico

O artigo científico é um dos principais instrumentos do diálogo entre pesquisadores. É parte de uma publicação com autoria declarada, que apresenta e discute ideias, métodos, técnicas, processos e resultados nas diversas áreas do conhecimento (NBR 6022, 2003, p.2). Desta maneira, o seu objetivo é tornar público o conhecimento produzido permitindo o seu debate.

Diferente de outros trabalhos científicos, pela sua reduzida dimensão e conteúdo, os artigos científicos configuram pequenos estudos, completos, que tratam de um problema científico, de forma clara, concisa e fidedigna, mas que não se constituem em matéria de um livro (LAKATOS; MARCONI, 1985). Caracterizado por uma linguagem impessoal, técnica, objetiva e com características próprias de acordo com cada área da ciência, constrói-se através de métodos, configurando-se como um gênero do discurso acadêmico, dado o seu propósito comunicativo (BHATIA, 1993; SWALES, 1990). Entre os gêneros discursivos mais utilizados por pesquisadores, segundo pesquisa realizada por Motta-Roth (1998), estão os capítulos de livros e artigos de revistas científicas. O conceito de gênero não está vinculado especificamente à literatura: é utilizado para se referir a uma categoria de discurso de qualquer tipo, falado ou escrito, com ou sem aspirações literárias (SWALES, 1990).

Um gênero compreende uma classe de eventos comunicativos, cujos membros compartilham os mesmos propósitos comunicativos. Tais propósitos são reconhecidos pelos membros especialistas da comunidade discursiva de origem e, portanto, constituem o conjunto de razões (*rationale*) para o gênero. Essas razões moldam a estrutura esquemática do discurso e influenciam e impõem limites à escolha de conteúdo e de estilo. (SWALES, 1990, p.58)

Ao analisarmos a citação acima, dois aspectos devem ser considerados: o evento comunicativo e a comunidade discursiva.

O primeiro compreende “não somente o discurso e seus participantes, mas também o papel desse discurso e o ambiente de sua produção e recepção, incluindo suas associações

históricas e culturais” (SWALES, 1990), com características que os tornam compreensíveis para uma dada comunidade.

Para Swales (1990), o conceito de comunidade discursiva, segundo tópico a ser abordado, está relacionado com as pessoas que trabalham com um determinado gênero e que possuem conhecimento de suas convenções. Dominar tais convenções discursivas é condição necessária para alguém pertencê-la. Assim, o artigo científico se estabelece como um evento comunicativo, com propósitos compartilhados, dentro de um processo social, onde são lidos, avaliados por outros pesquisadores, citados, até que as novas descobertas neles descritos sejam finalmente incorporadas à ciência. Isso conduz a um forte aspecto social para ciência, pois está relacionado ao sucesso de um pesquisador com a habilidade de convencer seus pares da qualidade de seu trabalho e a validade dos seus argumentos. Um longo processo social de mediação, através da leitura do texto, onde é avaliado, criticado e citado pelos pares para que o novo conhecimento possa enfim ser incorporado ao acervo de conhecimento público aceito numa determinada ciência. De maneira resumida, pode-se dizer que a comunicação na visão de Swales (1990) se dá através do gênero que possuirá as características da comunidade (grupo social, contexto social) do qual se originou.

Como um evento comunicativo - que passa por transformações acompanhando as modificações nas formas de comunicação de uma sociedade - o artigo científico atravessou diferentes estágios, estilos e formas desde o seu aparecimento, na Inglaterra, na *Royal Academy*, o *Philosophical Transactions* (MEADOWS, 1974; GARVEY, 1979 ZIMAN, 1979) até adquirir suas formas atuais.

Essas transformações podem ser observadas através de estudos históricos como salienta Moraes (2005, p.61) indicam “uma mudança da descrição para a explanação”, o que “sugere um aumento na complexidade intelectual”, além que “houve uma diminuição no uso de verbos de relatar (ex: *Smith reports...*) e aumento dos verbos de ação (ex: *Temperature increases ...*)” o que também sugere “que a descoberta ou teoria foi trazida para uma posição gramatical central, enquanto os cientistas passaram a ter menor destaque, acrescentando, assim, densidade à discussão e integrando o material de fonte à continuidade do argumento”. O texto passa, portanto, a ser um registro de escrita mais impessoal, “centrado no autor”, “centrado no objeto” de estudo (ATKINSON, 1999).

2.2 A Organização do Artigo Científico

Para organizar e estruturar o fluxo de informação no texto é comum utilizar o formato Introdução Corpo Conclusão. Este formato narrativo persiste por mais de dois mil anos por

sua flexibilidade, escalabilidade, tornando-se uma poderosa ferramenta por proporcionar para os leitores a informação que eles precisam quando eles dela necessitam.

Na introdução, encontra-se o objetivo, assunto ou problema, a ideia defendida pelo o autor, formuladas através de proposições. No desenvolvimento do texto, busca-se fundamentar e analisar a tese e apresentar os argumentos estabelecendo relações de causa e efeito ou comparações entre situações, fazendo-se usos de citações de outros autores, dados estatísticos e experiências. Uma conclusão ou discussão finaliza o texto, onde a tese e uma síntese das ideias são reapresentadas propondo a solução do problema inicialmente exposto. Nesta parte do texto, é ainda possível a apresentação de novos elementos que surgiram no decorrer da experimentação das ideias, que podem se constituir em novas hipóteses para um futuro trabalho.

O que caracteriza o desenvolvimento do texto é que os parágrafos estão todos, ou deveriam estar relacionados com a tese ou ideia principal do autor. Não formam, pois, um conjunto de frases soltas e sem sentido. O texto é organizado de maneira intencionada a convencer os interlocutores sobre as proposições e argumentações do autor. Busca-se uma coerência textual, isto é, os diversos parágrafos devem estar encadeados de forma lógica para garantir a argumentação.

De maneira geral, a coesão é a associação consistente dos elementos textuais ou modo como eles são ligados, enquanto a coerência é a ligação do conjunto de elementos formativos de um texto, estando diretamente ligados aos conceitos e às relações semânticas, que permitem a união dos elementos textuais; ela não é independente do contexto no qual o texto está inscrito (BEAUGRANDE; DRESSLER, 1987). Assim, o artigo científico organiza-se através de uma estrutura dissertativo-argumentativo, onde são discutidos os problemas, as hipóteses, uma descrição dos métodos utilizados para obter os resultados (com a sua posterior discussão ou conclusão), buscando persuadir seus pares com relação a seu ponto de vista. A persuasão se dá não no sentido da sedução, mas é construída através de uma linha de raciocínio lógico, expressa através de uma coerência textual, onde a sua consistência e a apresentação de provas (fatos, dados estatísticos, experimentos, etc.) são bases da argumentação (GARCIA, 1992). Os cientistas se esforçam para convencer seus pares da validade de suas descobertas dentro de um contexto persuasivo tornando o artigo científico um processo retórico (GROSS, 1990).

O esquema utilizado para organizar o texto possivelmente não é arbitrário e está relacionado com os aspectos semânticos e pragmáticos. Assim, reflete determinadas funções cognitivas, pragmáticas ou sociais da comunicação através do texto (VAN DIJK, 1978).

Gradualmente, ao longo século XX, a estrutura IMRAD (introdução, métodos, resultados e discussão), passa a ser adotada, sobretudo, nas áreas médicas. IMRAD é uma variação do IBC e, segundo Sollaci e Perreira (2004), essa estrutura já era considerada ideal para os textos científicos no início do século XX, porém não era usada pelos autores em geral. Ainda segundo os autores, após a segunda guerra, conferências internacionais em publicação científica recomendaram este formato culminando com as diretrizes fixadas pelo *International Committee of Medical Journal Editors*.

O texto de artigos observacionais e experimentais é normalmente (mas não necessariamente) dividido em seções com a Introdução Métodos, Resultados, e Discussão. Esta estrutura denominada de "IMRAD" simplesmente não é um formato de publicação arbitrário, porém uma reflexão direta do processo de descoberta científica. Artigos longos podem precisar de subtítulos dentro de algumas seções (especialmente os Resultados e Discussão) para deixar claro o conteúdo. Outros tipos de artigos, como caso relatórios, revisões, e editoriais, provavelmente necessitam de outros formatos. (INTERNATIONAL COMMITTEE OF MEDICAL JOURNAL EDITORS)³⁸

Atualmente a estrutura IMRAD é uma das mais contempladas em artigos científicos, na área das ciências médicas, o que necessariamente não ocorre em outras áreas. Foi influenciada pelos editores das revistas científicas cuja padronização torna a revisão, pelos pares, mais fácil (HUTH, 1987).

Segundo Curty e Boccatto (2005), a normalização técnica dos artigos que compõem um fascículo de um periódico é um dos atributos que permite avaliar a qualidade da publicação e não pode ser encarada como limitador da criatividade e liberdade dos autores.

No Brasil, a busca pela padronização, também, pode ser observada pela aplicação das normas da Associação Brasileira de Normas Técnicas (ABNT). Sem dúvida, a adoção de padronização como IMRAD e NBR são consequências do crescimento constante da comunicação científica (MEADOWS, 1999). Tal estrutura, além de facilitar a leitura, a compreensão do texto e a revisão dos pares, facilita, também, a execução da análise documental que visa representar o conteúdo do documento sobre uma forma diferente do original para sua posterior consulta e referência (BARDIN, 1977).

2.3 O Texto do Artigo Científico

Segundo Van Dijk (1978), um texto apresenta três estruturas constituintes. A primeira, microestrutura, situa-se no nível de sentenças e orações que vão satisfazer as condições de coerência e conexão. Uma segunda, a macroestrutura, refere-se à ideia geral do texto, à

³⁸ Disponível em: <http://www.icmje.org/>

representação abstrata do seu significado (a estrutura semântica). E, finalmente, a superestrutura, que caracteriza o tipo de texto, sua forma, suas unidades, a organização sequencial de suas partes.

Os artigos científicos têm por objetivo relatar resultados de pesquisa e sua confecção é influenciada pela comunidade discursiva que estabelece convenções comunicativas e estruturais, tal como IMRAD. O formato IMRAD é uma estrutura esquemática global que um texto científico possui e elas desempenham um papel importante no processo de produção, compreensão e condicionamento do discurso (VAN DIJK, 1978). Essas estruturas (esquemas) além de serem responsáveis pela organização do texto, pois definem regras e categorias, algumas obrigatórias, outras opcionais, determinam o gênero e as características culturais da comunidade discursiva.

Segundo Swales (1990), a construção da argumentação é refinada em movimentos argumentativos não hierárquicos, isto é, trechos retoricamente coerentes de texto. A construção dos trechos conta com o auxílio do IMRAD, que define a estrutura argumentativa formada por categorias pré-definidas como introdução, materiais e métodos, entre outros. Assim, de acordo com cada categoria da superestrutura, o autor escolhe palavras e sentenças para sua composição textual, buscando a coerência do texto e o convencimento do leitor através da argumentação lógica. O autor utiliza, também, de sinais relevantes e “pistas textuais”. As pistas textuais se apresentam como marcas gráficas (tipo de letra; letra sublinhada; caixa alta), marcas sintáticas (expressas tipos de palavras), além de marcas lexicais (expressas por palavras e expressões, tais como: "o objetivo é", "a conclusão é") para sinalizar, junto ao leitor, sua argumentação.

Embora as estruturas sejam independente uma da outra elas estão intimamente ligadas. A superestrutura organiza a macroestrutura dando sequenciamento e ordenamento ao conteúdo do texto ao impor certas regras e limitações (KINTSCH; VAN DIJK, 1983). A superestrutura sustenta o assunto, a ideia, descrita no texto e no caso do artigo científico, a estrutura dissertativo-argumentativo.

A análise dos textos dos artigos científicos sob a ótica do propósito comunicativo estabelece-se como um instrumento para entender a própria noção de gênero e, por outro lado, cria condições para que estratégias pedagógicas possam ser definidas para ensinar estudantes de pós-graduação e acadêmicos a escrever de forma clara seus artigos.

O trabalho de Swales (1990) é marcante neste contexto, apresentando o *CARS model* (*Create a Research Space*), com suporte em duas unidades de análise: os movimentos (*move*) e os passos (*step*). O movimento é uma unidade semântica relevante para o propósito do autor,

isto é, “um segmento de texto que compõe um pacote de características linguísticas (significados léxicos, significados proposicional, a força ilocutória, etc.) o qual [dá] o segmento uma orientação uniforme e sinaliza o conteúdo de discurso” (NWOGU, 1997, p. 114) e os passos são subobjetivos com os quais os movimentos são construídos. Swales utiliza o *CARS model* para obter uma explicação de como a estrutura de um gênero se constitui.

2.3.1 Os Movimentos Comunicativos na Seção de Introdução

Ao analisar a introdução dos artigos científicos Swales (1990, p.141) apresenta um conjunto de movimentos e passos apontando que um dos passos utilizados pelo autor é delinear os objetivos de seu trabalho. Já Skelton (1994) analisou 50 artigos publicados no *British Journal of General Practice* e descrevendo a estrutura comunicativa do texto. Segundo o autor, a última função encontrada na introdução aponta para o objetivo do trabalho onde são utilizadas palavras indicadoras tais como: *aim, intention, purpose*, etc. O trabalho de Nwogu (1997) baseia-se em artigos da área médica e, também, na introdução do artigo a referência para o objetivo da pesquisa é apontada. Kanoksilapatham (2005) afirma que todos os movimentos identificados na seção de introdução de artigos científicos na área de bioquímica são convencionais e entre eles a declaração do objetivo.

Como pode se observar esses trabalhos apontam que a introdução de um artigo científico é um dos locais onde possivelmente será encontrado o objetivo do trabalho.

2.3.2 Os Movimentos Comunicativos no Abstract

A norma ISO 214 define o abstract como uma “representação abreviada e exata do conteúdo de um documento” que possibilita ao leitor “identificar rapidamente e acertadamente o conteúdo básico do documento para determinar sua relevância”. A norma estabelece a seguinte estrutura para os abstracts: 1) **Background**, a indicação do contexto da pesquisa, a pesquisa prévia e teorias subjacentes; 2) **Propósito** ou escopo do trabalho, orações que descrevem razões que motivaram estudo e metas; 3) **Hipóteses**, orações incluindo afirmações que podem ser aceitas ou rejeitadas com base nos resultados apresentados; 4) **Método**, descrição do processo de pesquisa; 5) **Resultados** teóricos ou empíricos, dados, relações, efeitos, influências; e 6) **Conclusões** e discussão, orações que descrevem as implicações dos resultados (normalmente conectados com as metas do estudo, argumentos por aceitar ou rejeitar a hipótese e sugestões para futura pesquisa).

Tibbo (1993) analisou 30 *abstracts* de artigos na área de Química, Psicologia e História observando a existência de elementos diferentes daqueles posposto pela ISO 214 e a baixa frequência de ocorrência da hipótese. Não há uma concordância uniforme com relação à estrutura de movimentos (PEACOCK, 2002) e embora o discurso científico esteja cada vez mais internacionalizado, ele é fortemente direcionado pela área da ciência e pela cultura dos autores, que, mesmo escrevendo na língua inglesa “preservam seu estilo nativo” (GOLEBIOWSKI, 1999).

3 AS FRASES INDICATIVAS

É comum o autor, por conta da estrutura argumentativa formada pelo IMRAD e a necessidade de convencer seus pares, com relação à importância do trabalho, escolher palavras e sentenças para sua composição textual, utilizando “pistas textuais”, expressas por palavras e expressões, tais como: "o objetivo é", "a finalidade do trabalho é" (VAN DICK, 1992), de maneira a deixar claro qual o seu objetivo no trabalho. A utilização de frases indicativas não são comuns nos resumos estruturados, em que os textos se iniciam geralmente com verbos no infinitivo (*infinitive tense*).

A fim de identificar, caracterizar e apontar em que locais da estrutura IMRAD as frases indicativas do objetivo ocorrem, procedeu-se à análise dos setenta dois (72) artigos, apontando a existência de oito artigos (11,11%) que não apresentavam frases que indicavam claramente o objetivo. Foram identificadas 92 frases no total (em 64 artigos), sendo que 29,03% ocorreram mais de uma vez no corpo do artigo.

Os locais onde as frases ocorreram mostraram um equilíbrio: 44,57% ocorrem no resumo (abstract), enquanto outros 43,48% no final da introdução, confirmando os resultados apontados por outros trabalhos na área de movimentos comunicativos (NWOGU, 1997; SKELTON, 1994). Foi possível identificar a utilização de alguns substantivos, nas frases indicativas, tais como: estudo, relatório, trabalho, artigo, investigação, propósito, objetivo e a ocorrência de verbos associados, aos seguintes tempos verbais (língua Inglesa): *present tense*, *past tense*, *present perfect tense*, *passive voice*. Foi observada, também, uma tendência dos autores a utilizarem o mesmo padrão linguístico, dentro do mesmo periódico.

Essa característica marcante na forma de o autor anunciar o seu objetivo possibilita a construção de modelos de regras ou gramáticas que podem ser facilmente implementadas, como parte de uma estratégia de processamento automático do texto.

processo de seleção. Apenas as frases que possuem *Our, Here, In, The, This, We* ou contém algum termo de um dos dicionários sofrem rotulação gramatical.

5 EXPERIMENTOS

Para testar a eficácia do modelo construído para a extração do objetivo do artigo científico, foi construída uma aplicação WEB que implementa o modelo e que apoiou a realização de diferentes experimentos.

5.1 Primeiro Experimento

Foram selecionados, aleatoriamente, 50 artigos classificados como contribuições originais, do periódico JAMA (*Journal of the American Medical Association*). O JAMA possui um resumo estruturado, que apresenta informações claras, diretas e identificáveis, através de rótulos que caracterizam o objetivo do trabalho, o contexto da pesquisa, a metodologia usada, resultados obtidos e a conclusão. A escolha de um periódico com resumo estruturado teve como objetivo eliminar a interpretação e escolha do objetivo, já que o próprio autor indica-o. Para cada artigo selecionado foram retirados os rótulos que identificavam as informações no resumo estruturado, mantendo-se apenas o conteúdo do resumo. Em seguida, todo o texto dos artigos (sem as marcações do resumo estruturado) foi submetido ao programa que extraiu as frases consideradas como objetivo do artigo. O programa extraiu frases em 88% (44) dos artigos e em 12% (6) não forneceu nenhuma informação.

De posse das respostas obtidas sobre os 44 artigos, seis avaliadores verificaram se a frase extraída era semelhante com a frase definida no resumo estruturado, pelo autor, como sendo o objetivo do artigo. Cada julgador votava se concordava ou não com a semelhança. O comitê foi formado por quatro professores universitários (um da Ciência da Computação, um da Farmacologia e dois da Ciência da Informação) e um aluno da área de biologia; cada avaliação foi realizada individualmente. Por seis vezes (13,6%) houve a unanimidade em apontar que a frase extraída não era o objetivo e, por dez vezes (22,7%), a unanimidade em apontar que a frase era o objetivo, enquanto nos outros 63% (28) os julgadores divergiram de opinião. Levando em consideração todos os artigos inclusive os que não tiveram a indicação do objetivo os percentuais são menores.

TABELA 1 – Avaliação dos julgadores

	Total Concordância (44)	Total Concordância (50)
Era frase extraída é o objetivo	22,7% (10)	20% (10)

Era frase extraída não é o objetivo	13,6% (6)	12%(6)
-------------------------------------	-----------	--------

Fonte: Próprio autor

Para análise da concordância interavaliadores, foi utilizado o índice estatístico *Kappa Fleiss*. O índice estabelece um grau de concordância, de acordo com número fixo de julgadores, quando eles classificam itens.

O valor do índice *Kappa Fleiss*, obtido para a avaliação dos 5 julgadores, 44 itens e duas categorias (sim e não) foi de 0,40321857, que pode ser interpretado como uma concordância justa (“apertada”). Esse valor foi considerado baixo e duas possíveis explicações são: o grupo de julgadores escolhido foi bastante heterogêneo, alguns deles não tinham conhecimento na área de biomedicina, valendo-se, apenas para seu julgamento da forma como as frases foram sintaticamente construídas; e o tipo de artigo escolhido (calcado no resumo estruturado) que “foge” das características do artigo clássico em biomedicina, isso porque o resumo apresenta um rótulo objetivo no qual o autor anuncia diretamente aquilo que irá realizar. Essa característica faz com que o autor não necessite mencionar no corpo do artigo qual é o seu objetivo e, conseqüentemente, não utiliza frases indicativas para tal, tornando o processo de extração baseado no modelo de regras ineficiente.

5.2 Segundo Experimento

TABELA 2 – Resultado da avaliação dos pesquisadores - Intervalo Confiança a 95% para a média da proporção

	Precisão	Revocação	<i>F-measure</i>
1	33,33%	100,00%	50,00%
2	100,00%	100,00%	100,00%
3	100,00%	100,00%	100,00%
4	0,00%	0,00%	0,00%
5	100,00%	100,00%	100,00%
6	25,00%	100,00%	40,00%
7	40,00%	100,00%	57,14%
8	100,00%	100,00%	100,00%
9	100,00%	100,00%	100,00%
10	20,00%	100,00%	33,33%
Média	61,83%	90,00%	68,05%
Desvio Padrão	0,42	0,32	0,37
Limite inferior	32,42%	67,59%	41,98%
Limite superior	91,25%	100,00%	94,11%

Fonte: Próprio autor.

A fim de eliminar a falta de conhecimento na área de biomedicina pelos avaliadores, foram convidados 10 professores, pesquisadores do Depto. de Farmacologia e Fisiologia da Universidade Federal Fluminense, que submeteram seus trabalhos, já publicados em diferentes revistas internacionais, ao programa de extração do objetivo que selecionou as frases que indicativas do objetivo.

TABELA 3 – Resumo da avaliação dos pesquisadores - Intervalo Confiança a 95% para a média da proporção

	Média Geral	Limite inferior	Limite superior
Revocação	90,00%	67,59%	100,00%
Precisão	61,83%	32,42%	91,25%
F-measure	68,05%	41,98%	94,11%

Fonte: Próprio autor.

Em seguida, cada autor julgou se uma das frases selecionadas seria o objetivo de seu trabalho; os resultados obtidos estão descritos nas tabelas anteriores. A precisão obtida (61,83%) foi prejudicada principalmente pela ambiguidade encontrada em algumas frases, como por exemplo: “*we analysed*”.

5.3 Terceiro Experimento

Buscando melhorar a precisão do modelo, novos termos foram inseridos nos dicionários e foi projetada uma nova avaliação. Foram selecionados 20 artigos do periódico Memórias do Instituto Oswaldo Cruz, 20 do *Brazilian Journal of Medical and Biological Research* e 20 do *Nature*, totalizando 60 artigos (tabela 4). Os resultados obtidos (tabela 5) foram de 94,85% para a revocação e 82,95% para a precisão, com uma exatidão de 86,6%. A diferença 11,90% é compatível com o valor anteriormente encontrado (11,11%) no estudo preliminar (COSTA, 2008) em que artigos não apresentaram claramente frases indicativas de seu objetivo.

Analisando cada uma das características individualmente obtemos os seguintes resultados e conclusões:

5.3.1 Revocação

Observa-se que não há diferença estatisticamente significativa entre a revocação nesses três periódicos. O teste de *Kruskal-Wallis*, ao nível de significância $\alpha = 0,05$, indica inexistência de diferença estatisticamente significativa ($p > 0,05$) entre os três periódicos quanto à revocação medida percentualmente ($H = 2,005$; g.l. = 2; valor- $p = 0,367$).

5.3.2 Precisão

O teste de *Kruskal-Wallis*, ao nível de significância $\alpha = 0,05$, indica inexistência de diferença estatisticamente significativa ($p > 0,05$) entre os três periódicos quanto à Precisão medida percentualmente ($H = 3,094$; g.l. = 2; valor-p = 0,213).

5.3.3 F-measure

O teste de *Kruskal-Wallis*, ao nível de significância $\alpha = 0,05$, indica inexistência de diferença estatisticamente significativa ($p > 0,05$) entre os três periódicos quanto à F-measure medida percentualmente ($H = 3,913$; g.l. = 2; valor-p = 0,141).

TABELA 4 – Resultados obtidos para a revocação e a precisão - Intervalo Confiança a 95% para a média da proporção

	<i>Nature</i>			Memórias do Instituto Oswaldo Cruz			<i>Brazilian Journal of Medical and Biological</i>		
	Rev.	Prec.	F	Rev.	Prec.	F	Rev.	Prec.	F
1	100,00	100,00	100,00	100,00	66,00	79,52	100,00	66,00	79,52
2	100,00	100,00	100,00	100,00	100,00	100,00	100,00	66,00	79,52
3	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
4	100,00	100,00	100,00	100,00	50,00	66,67	50,00	100,00	66,67
5	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
6	100,00	50,00	66,67	100,00	100,00	100,00	100,00	50,00	66,67
7	100,00	100,00	100,00	100,00	100,00	100,00	100,00	75,00	85,71
8	100,00	100,00	100,00	100,00	100,00	100,00	100,00	50,00	66,67
9	100,00	100,00	100,00	100,00	100,00	100,00	100,00	50,00	66,67
10	50,00	100,00	66,67	100,00	50,00	66,67	100,00	100,00	100,00
11	100,00	100,00	100,00	100,00	100,00	100,00	100,00	50,00	66,67
12	100,00	33,00	49,62	100,00	100,00	100,00	66,00	66,00	66,00
13	100,00	66,00	79,52	100,00	100,00	100,00	100,00	100,00	100,00
14	50,00	50,00	50,00	100,00	66,00	79,52	100,00	100,00	100,00
15	100,00	100,00	100,00	50,00	50,00	50,00	100,00	100,00	100,00
16	100,00	100,00	100,00	100,00	100,00	100,00	75,00	75,00	75,00
17	100,00	100,00	100,00	100,00	100,00	100,00	50,00	50,00	50,00
18	100,00	100,00	100,00	100,00	50,00	66,67	100,00	66,00	79,52
19	100,00	100,00	100,00	100,00	66,00	79,52	100,00	66,00	79,52
20	100,00	50,00	66,67	100,00	100,00	100,00	100,00	100,00	100,00
Média	95,00	87,45	88,96	97,50	84,90	89,43	92,05	76,50	81,41
Des. padrão	15,39	22,94	18,28	11,18	21,65	15,99	17,04	21,13	15,90
Limite inf.	88,26	77,40	80,94	92,60	75,41	82,42	84,58	67,24	74,44
Limite sup.	100,00	97,50	96,97	100,00	94,39	96,43	99,52	85,76	88,37

Fonte: Próprio autor.

TABELA 5 – Resultados obtidos para a revocação e a precisão - Intervalo Confiança a 95% para a média da proporção

	Média Geral	Limite inferior	Limite superior
Revocação	94,85	91,14	98,56
Precisão	82,95	89,27	100,00
<i>F-measure</i>	86,60	82,32	90,87

Fonte: Próprio autor.

A precisão obtida (82,95%) aumentou com relação ao experimento anterior, porém continuou sendo prejudicada, principalmente, pela ambiguidade encontrada em algumas frases (por exemplo: “*we analysed*”).

6 CONCLUSÃO

O objetivo é um dos elementos fundamentais no trabalho científico sendo um dos elementos de OCCA. A análise do objetivo buscou identificar quais as estruturas sintáticas mais utilizadas nos artigos. Essa análise propiciou a elaboração de um modelo de regras que foi utilizado no processo de extração automática do objetivo no texto do artigo científico.

O modelo regra foi sendo melhorado a cada experimento sendo obtido uma precisão de 82,95% no último experimento, porém continuou sendo prejudicada, principalmente, pela ambiguidade encontrada em algumas frases (por exemplo: “*we analysed*”). Esse tipo de frase, como já comentado, por ser usada tanto para indicar o objetivo geral do artigo quanto para indicar um passo (ou atividade) dentro da pesquisa, nesse caso, ela não deixa de ser um objetivo, porém secundário (subobjetivo).

O estudo e experimentos aqui relatados são o ponto de partida para a busca de outras soluções que possam contribuir para que a precisão na extração automática possa ser aumentada.

REFERÊNCIAS

- ATKINSON D. **Scientific discourse in sociohistorical context**: the Philosophical Transactions of the Royal Society of London, 1675–1975. Mahwah, NJ: Lawrence Erlbaum, 1999.
- BARDIN, L. **Análise de Conteúdo**. Lisboa, Portugal: Edições 70, 1970.
- BEAUGRANDE, R.; DRESSLER, M. U. Introduction to text linguistics. In: Congress of Linguists, 14, 1987, Berlin. **Proceedings...** Berlin, 1987. Disponível em: http://www.beaugrande.com/introduction_to_text_linguistics.htm. Acesso em: 31 jul. 2008.
- BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The semantic web. **Scientific American Magazine**, 2001, p. 29-37. Disponível em: <http://www.scian.com/2001/0501issue/0501/berners-lee.html>. Acesso em: 24 mai. 2001.

BHATIA, V. K. **Analysing genre: language use in professional settings**. London: Longman, 1993.

BUNGE, M. **Philosophy of science**. New Brunswick, London: Transaction Publishers, 1998. 2v.

COSTA, Leonardo Cruz da. Uma proposta de processo de submissão de artigos científicos às publicações eletrônicas semânticas em ciências biomédicas. In: ASOCIACIÓN DE EDUCACIÓN E INVESTIGACIÓN EN CIENCIA DE LA INFORMACIÓN DE IBEROAMÉRICA Y EL CARIBE, 2011, Marília. **Anais...** Marília: Edicic. Disponível em: <<http://www.edicic.org/revista>>. Acesso em: 18 jul. 2014.

COSTA, Leonardo Cruz da; MARCONDES, C. H. Padrões linguísticos para identificar elementos de ontologia. In SEMINÁRIO DE PESQUISA EM ONTOLOGIA NO BRASIL, 2008, Niterói. **Anais...** Niterói: UFF.

CURTY, M. G.; BOCCATO, V. R. C. O artigo científico como forma de comunicação do conhecimento na área de Ciência da Informação. **Perspectivas em Ciência da Informação**, Belo Horizonte, v.10, n.1, p. 94-107, jan./jun. 2005.

GARCIA, O. M. **Comunicação em prosa moderna**. 15.ed. Rio de Janeiro: Fundação Getúlio Vargas, 1992.

GARVEY, W. D. **Communication: essence of science; facilitating information exchange among librarians, scientists, engineers and students**. Oxford: Pergamon Press, 1979.

GOLEBIEWSKI, Zosia. Application of Swales' model in research papers by non-native authors. **International Review Of Applied Linguistics**. Heidelberg, vol.37 n.3, p. 231-248. mar. 1999.

GROSS, A. G. **The rhetoric of science**. Cambridge: Harvard University Press, 2006.

HUTH, EDWARD J. Needed: Review Articles with More Scientific Rigor. **Annals Internal Medicine**. v.106, p. 470-471, 1987.

KANOKSILAPATHAM, Budsaba. Rhetorical structure of biochemistry research articles. **English For Specific Purposes**. Philadelphia, v.24, n.3, p. 269-292, mar. 2005.

KINTSCH, W. ; VAN DIJK, T. A. **Strategies of discourse comprehension**. San Diego, California, Academic Press, 1983.

LAKATOS, E. M.; MARCONI, M. A.. **Fundamentos de metodologia científica**. São Paulo: Atlas, 1985.

MALHEIROS, L.; MARCONDES, C. H. A identificação de novas descobertas científicas através da análise do conhecimento contido em artigos científicos. In: ENCONTRO DA ASSOCIAÇÃO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 7., 2006, Marília. **Anais...** Marília, ENANCIB, nov. 2006. 1 CD-ROM.

MARCONDES, Carlos Henrique. From scientific communication to public knowledge: the scientific article Web published as a knowledge base. In: **ELPub – INTERNATIONAL CONFERENCE ON ELECTRONIC PUBLISHING**, 9., 2005, Leuven, Bélgica. Proceedings...Leuven: Peeters Publishing, 2005. p.119-27.

MARCONDES, Carlos Henrique.; MENDONÇA, Marília Alvarenga Rocha; MALHEIROS, Luciana Reis; COSTA, Leonardo Cruz da; SANTOS, Tatiana. Cristina Paredes. Bases ontológicas e conceituais para um modelo do conhecimento científico em artigos biomédicos. **Revista Eletrônica de Comunicação, Informação & Inovação e Saúde**. v.3, n.1, 2009, p.19-30.

MEADOWS, A. J. **A comunicação científica**. Brasília, Briquet de Lemos, 1999.

MEADOWS, A. J. **Communication in Science**. London: Butterworths, 1974. 248p.

MILLER, D.L., Explanation versus description. **Philosophical Review**, v.56, n.3, p.306-312, 1947.

MORAES, L. S. B. **O Metadiscorso em artigos acadêmicos: variação intercultural, interdisciplinar e retórica**. 2005. Tese (Doutorado) - Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2005.

MOTTA-ROTH, D. (Org.). **Leitura em língua estrangeira na escola: teoria e prática**. Santa Maria: Universidade Federal de Santa Catarina, 1998.

NWOGU, K.N. The medical research paper: structure and functions. **English for Specific purposes**, v.16 n.2, p.28-32, 1997.

PASSIN, T. B. **Explorer's guide to the semantic web**. United States of America: Manning Publications Co.Greenwich, California. 2000.

PEACOCK, M. Communicative moves in the discussion section of research articles. **System**, v.30, n.4, p.479-97, 2002.

SKELTON, J. R.; EDWARDS, S. J L. The function of the discussion section in academic medical writing. *British Medical Journal*, v.320, p. 1269-1270, 2000.

SOLLACI, Luciana B.; PEREIRA, Mauricio G. The introduction, methods, results, and discussion (IMRAD) structure: a fifty-year survey. **Journal of the Medical Library Association**, Chicago, v. 92, n. 3, p.364-371, abr. 2004.

SWALES, J. M. **Genre analysis: english in academic and research settings**. Nova Iorque: Cambridge University Press, 1990.

TIBBO, H.R. **Abstracting, information retrieval, and the humanities: Providing access to historical literature**. Chicago, London: American Library Association, 1993.

VAN DIJK, T. **Cognição, discurso e interação**. São Paulo: Contexto, 1992.

VAN DIJK, T. **La ciência del texto: um enfoque interdisciplinar**. Tradução de Sibila Hunzinger. Barcelona: Paidós, 1978.

ZIMAN, J. **Conhecimento público**. Belo Horizonte: Itatiaia, São Paulo: Editora da Universidade de São Paulo, 1979.

PERSONALIDADE E MATÉRIA NA TEORIA DA CLASSIFICAÇÃO FACETADA: A QUESTÃO DO CONTEXTO, PRESSUPOSTOS TEÓRICOS E METODOLÓGICOS

PERSONALITY AND MATTER IN FACETED CLASSIFICATION THEORY: ISSUES ON CONTEXT AND THEORETICAL AND METHODOLOGICAL BACKGROUND

Linair Maria Campos
Maria Luiza de Almeida Campos

Resumo: A Teoria da Classificação Facetada, de Ranganathan, tem sido utilizada recentemente como base para a organização de domínios de conhecimento nas mais diversas áreas, uma vez que fornece um conjunto de categorias de alto nível que possibilitam um princípio norteador para o recorte desses domínios. Originalmente, entretanto, seu foco de uso era voltado para a organização de documentos nas estantes e para a descrição de assuntos em bibliotecas. Essa mudança de propósitos, bem como a dificuldade de entendimento da obra de Ranganathan, parece trazer um custo associado para as iniciativas de aplicação de suas teorias, em especial no que tange à identificação das categorias de Personalidade e Matéria. Corroborando essa proposição, temos que a análise de alguns exemplos de classificação presentes nas obras de referência de Ranganathan, quando confrontadas com as definições apresentadas para as suas categorias revelam certas divergências. O objetivo do presente artigo é apresentar uma análise dessas divergências, ilustrando a importância do contexto e do entendimento dos pressupostos teóricos na utilização de categorias para fins de organização de conhecimento. Como resultado apresenta-se uma sistematização de questões suscitadas a partir dessa análise, que denominamos de análise ontológica, e mostra-se a importância de se considerar o contexto na classificação facetada, bem como de entender os pressupostos que norteiam a definição de categorias que apoiam uma estrutura classificatória.

Palavras-chave: Análise Ontológica, Teoria da Classificação Facetada, Contexto.

Abstract: Ranganathan's Faceted Classification Theory has recently been used as a basis for organizing domains of knowledge in several areas, as it provides a set of high-level categories that allow a guiding principle to represent these areas. Originally, however, his focus usage was the organization of documents on the shelves and to describe subjects in libraries. This change of purpose as well as the difficulty in understanding the work of Ranganathan seems to bring a cost associated to the initiatives of application of his theories, in particular regarding the identification of categories of Personality and Matter. Supporting this proposition, we have the analysis of some examples of classification in Ranganathan's work, when confronted with the current settings to their categories reveal certain differences. The objective of this paper is to present an analysis of these differences, illustrating the importance of context and understanding of the theoretical use of categories for the purpose of organizing knowledge assumptions. As a result, we present a systematization of questions arising from this analysis, we call ontological analysis, and shows the importance of considering the context in faceted classification as well as to understand the assumptions that guide the definition of categories that support a classificatory structure.

Keywords: Ontological Analysis, Faceted Classification Theory, Context.

1 INTRODUÇÃO

Nas atividades do profissional de informação, áreas de conhecimento são representadas, através de estruturas classificatórias, que vem através dos tempos se adequando

a diversas Teorias de Representação, que por vezes não são explicitadas por seus criadores. Estas teorias se colocam no espaço da elaboração de Modelos Conceituais independentes de conteúdos específicos de domínios. São metas representações utilizadas para a organização de conhecimento em domínios diferenciados. Tais princípios, com advento da Web Semântica e de ferramentas semânticas como as ontologias são de valor inestimável para a elaboração de modelos conceituais consistentes, os quais reforçam a ideia da necessidade de um perfil analítico na formação do profissional da informação.

No contexto da biblioteconomia, uma Teoria de Representação bastante estudada e que ainda hoje é base para a organização de conteúdos informativos, já em um espaço onde imperam as novas tecnologias (CAMPOS, GOMES e OLIVEIRA, 2013; GIESS, WILD e McMAHON, 2008), é a Teoria de Classificação Facetada de Ranganathan (1967a).

A Teoria da Classificação Facetada foi desenvolvida por Ranganathan, na década de 1930, para a organização de material bibliográfico. A novidade introduzida por ele na criação de sua tabela a *Colon Classification* foi a ruptura com uma tradição de construção de tabelas de classificação. Até então, elas descreviam um assunto conforme dispostos nas disciplinas ou nos grandes tratados. As tabelas desenvolvidas deste modo dificultavam, ou até mesmo impediam a representação de temas novos. A Tabela de Classificação produzida (*Colon Classification*) permitiu mapear uma área de assunto independentemente de como este se encontra estruturado na literatura. Qualquer assunto já estabelecido ou mesmo novo pode ser representado, pois existem regras para isto. Ranganathan procurou organizar, sistematizar, idéias, conceitos, fornecendo diretrizes para a sua categorização. Categorizar uma área de assunto é identificar nela os aspectos pelos quais tal assunto pode ser abordado, sendo a adoção do princípio da categorização fundamental, pois as categorias possibilitam uma visão da área como um todo inter-relacionado.

Nesse contexto, a Teoria da Classificação Facetada, de Ranganathan, tem sido utilizada recentemente como base para a organização de domínios de conhecimento nas mais diversas áreas, uma vez que possibilita que classes de alto nível sejam criadas a partir de um princípio norteador para o recorte desses domínios. Esse princípio norteador possivelmente teve inspirações na lógica, embora Ranganathan não afirme isso explicitamente. Em contraste, destaca que elas não têm nada a ver com o seu emprego em Metafísica (RANGANATHAN, 1967a, p. 399).

Ranganathan não deixa claro qual a fundamentação que adota para a definição de cada categoria, embora saibamos que foi fruto de sua larga experiência na classificação de assuntos, e que foi influenciado por trabalhos de pesquisadores tais como Sayers, Bliss,

Dewey, Shera, dentre outros (LA BARRE, 2010). Essa falta de explicitação mais precisa das categorias, aliada a uma redação que dá margem, por vezes, a entendimentos ambíguos, tem levado a dificuldades na compreensão e aplicação de sua teoria. Essa dificuldade, inclusive, é reconhecida por Ranganathan, quando afirma que:

Existem ainda algumas áreas de dúvida na distinção entre manifestações das categorias fundamentais "Energia", "Matéria" e "Personalidade". Essas áreas difíceis não aparecem com muita frequência. Portanto, pode-se ir prosseguindo, apesar dessa dificuldade momentânea, e resolver essas questões no devido tempo à medida que a experiência aumenta (RANGANATHAN, 1967b, p.14). Tradução nossa.

O que se pretende no presente trabalho é fazer uma análise da categorização de conceitos através da Teoria da Classificação Facetada de Ranganathan, para fins de formação de assuntos visando à indexação de documentos, procurando apontar questões que levam a um entendimento ambíguo de suas categorias, com foco nas categorias de Personalidade e Matéria. Como resultado apresenta-se um quadro sistemático com questões a serem consideradas para essas categorias, evidenciando a importância do contexto e do compromisso ontológico assumido na elaboração de modelos do domínio. Por outro lado, estamos também preocupados, no escopo de nosso Grupo de Pesquisa, com a formação de "classificacionistas", como denominava Ranganathan, ou dito de outra forma, de modelizadores, que são aqueles que elaboram classificações e não somente as usam. Consideramos, assim, que a discussão de tais Teorias de Representação são fundamentais, e tais princípios são atualmente, com advento da Web semântica e de ferramentas semânticas, como as ontologias, de valor inestimável para a elaboração de modelos conceituais consistentes independentes de domínios.

O restante desse artigo está estruturado como se segue: na seção 2 faz-se a descrição das categorias de Ranganathan; na seção 3 serão descritas algumas divergências sobre a sua aplicação; na seção 4 será apresentada uma sistematização de questões associadas às categorias de Ranganathan; por fim, na seção 5 serão apresentados conclusões e trabalhos futuros.

2 AS CATEGORIAS DE RANGANATHAN E O CONTEXTO

Na área da Ciência da Informação (CI), no contexto das linguagens documentárias notacionais (e.g. tabelas de classificação) e linguagens documentárias verbais (e.g. tesouros), Ranganathan fornece apoio teórico para estruturação de domínios através da noção de categorias e da aplicação de uma série de cânones para organização de cadeias e renques.

Ranganathan também apresenta teorias sobre como formar assuntos, com o propósito de indexação e descrição bibliográfica. De acordo com Campos (2001), as teorias de Ranganathan se encontram apresentadas em quatro obras básicas, a saber: *Five Laws of Library Science*, *Prolegomena to Library Classification*, *Philosophy of Library Classification* e *Colon Classification*. Outra obra de Ranganathan que também faz menção às suas teorias é *Comunication and Classification* (RANGANATHAN, 1951).

A Teoria da Classificação Facetada parte da identificação das categorias nos domínios de conhecimento que se quer representar. Pressupõe-se uma análise do domínio em questão, a partir da identificação do assunto básico e de que categorias este assunto abarcará. A partir daí, dentro de cada domínio, este é analisado e seus elementos básicos identificados e agrupados em classes de acordo com características comuns. Essas classes são chamadas de facetas. Cada faceta, por sua vez, pertence a uma categoria fundamental.

Entende-se por categoria, no contexto desse artigo, os tipos mais gerais sob os quais se podem agrupar coisas, de natureza semelhante, destacando que não existe uma proposta consensual de quais categorias seriam as mais adequadas para cobrir todas as coisas existentes no mundo, sejam elas reais ou imaginárias, concretas ou abstratas. O estudo de um conjunto de tais categorias, suas características e relações, de fato, é objeto de estudo e debate entre filósofos de diferentes correntes (LOUX, 2006), e, de acordo com Mazzocchi e Gnoli (2010), serviram de inspiração para Ranganathan, que propõe um conjunto de cinco categorias, referenciadas pelo acrônimo PMEST: Personalidade, Matéria, Energia, Espaço e Tempo, para as quais ele fornece definições e alguns exemplos, como exposto a seguir.

Tempo é utilizado para designar noções usuais relativas a períodos de tempo, tais como milênio, século, ano e também as associadas à passagem do tempo tais como dia e noite, estações do ano. Essa categoria, ainda, inclui noções relacionadas à qualidade do tempo no âmbito da meteorologia, como, por exemplo, seco, chuvoso, tempestade. *Espaço* se refere a aspectos usuais de localização e também a limites de espaço, tais como a superfície da terra, ou ainda o espaço fora ou dentro desse limite. Objetos do domínio geográfico tais como continentes, países, oceanos e mares, desertos, florestas, montanhas e rios, também são manifestações dessa categoria. *Energia* se refere a ações de modo geral, podendo estas estar associadas a todos os tipos de entidades: inanimadas, animadas, conceituais, intelectuais e intuitivas. *Matéria* se refere à manifestação de materiais e propriedades intrínsecos às ideias e termos denotados no contexto da disciplina classificatória (RANGANATHAN, 1967a). Por exemplo, através da manifestação de *material*, uma mesa pode ser feita de madeira (que é um material), enquanto que através da manifestação de *propriedade* uma mesa pode ter uma

altura. Material e propriedade não fazem parte da mesa embora sejam intrínsecos à mesa, o que estaria de acordo com a ideia de Aristóteles sobre o aspecto da constituição, conforme discutido na seção 2.2. Além disso, tanto material quanto propriedades podem estar presentes em outras entidades. Ranganathan observa ainda que ideias isoladas tais como morfologia, fisiologia, ecologia, higiene e doença, dentre outras, que em um primeiro momento ele considerava como pertencentes à categoria de *Energia*, devem ser consideradas, na verdade, como manifestações da categoria de *Matéria* (RANGANATHAN, 1967a). Na verdade, na sétima edição de seu *Colon Classification* (RANGANATHAN, 1989), o autor propõe uma mudança no escopo da categoria de *Matéria*, afirmando que esta engloba não só os aspectos de material e propriedade, como também de método. Segundo Mazzocchi e Gnoli (2010), essa mudança é criticada por alguns autores, pois o aspecto de método poderia ser incluído na categoria de *Energia*, considerando-se que de fato, o conceito de método remete à ideia de ação, e se desdobra no tempo. Por fim, *Personalidade* se refere às entidades que não pertencem às outras categorias. (RANGANATHAN, 1967a), no que Ranganathan se refere como o *método dos resíduos*.

As categorias de Ranganathan são aplicadas de duas formas. Uma das aplicações se situa na organização de estruturas classificatórias, em domínios de conhecimento. Nesse sentido, as categorias fornecem um ponto de partida para uma abordagem *top-down* (de cima para baixo) para a formação das hierarquias de conceitos. Outra das aplicações é para orientar a indexação de documentos, quando da formação de assuntos. Lembrando-se que uma classificação sempre está sujeita a uma avaliação, a qual gera insumos para uma possível adequação na estrutura classificatória.

Para formar uma estrutura classificatória, as categorias fundamentais são divididas em facetas, e as facetas em isolados. Na *Colon Classification*, um isolado é uma unidade de conhecimento (SATIJA, 2002), a qual será utilizada para a composição de assuntos, no âmbito da classificação bibliográfica.

Ao utilizar-se a sua teoria para indexação de documentos, os conceitos que formam os assuntos destes são identificados com base nas categorias, seguindo uma ordem pré-determinada de composição, uma vez que uma obra pode ser indexada com mais de um conceito, de forma pós-coordenada.

2.1 O contexto na formação de uma estrutura classificatória

Ao definir as suas categorias fundamentais, Ranganathan não apresenta de forma separada a definição das categorias para a criação das estruturas classificatórias e a

identificação das categorias para a composição de assuntos. Dessa forma, a questão da influência do contexto no âmbito da formação da estrutura classificatória não é tão facilmente identificável em sua obra quanto o é no caso da composição de assuntos. De fato, Ranganathan fornece exemplos de assuntos categorizados sem indicar o seu contexto, como é o caso de higiene, doença, etc, dentre outros, que afirma pertencerem à categoria de Matéria (RANGANATHAN, 1967b). Entretanto, em sua obra *Communication and Classification* (1951), são apresentados exemplos em que a contextualização na elaboração de estruturas classificatórias é inequívoca, como no caso em que Personalidade representa conceitos relativos a tempo e espaço, nos domínios de Biografia, História e Direito:

374 Época cronológica

Em biografia, a categoria fundamental personalidade pode se manifestar como uma época cronológica, digamos, um ano ou uma década. É assim na Colon Classification, onde um biografado é representado pelo ano do seu nascimento. Na literatura, também, a Colon Classification representa um autor com um foco no tempo (RANGANATHAN, 1951, p. 375).

(...) 373 Área Geográfica

Em história e em direito, a categoria fundamental personalidade se manifesta como comunidade. Na história, uma comunidade é geralmente representada pela área geográfica ocupada por ela – um país ou grupo de países, um distrito, uma cidade, uma vila, ou qualquer combinação destes. Em direito, também, uma comunidade pode ser representada por uma área geográfica. (RANGANATHAN, 1951, p. 375).

Outro exemplo é mencionado no que tange ao conceito de substância. Na Física, uma substância como o ferro é considerada como Matéria, enquanto que na Química é considerada como Personalidade:

361 Matéria qua Matéria

Em primeiro lugar, temos de reconhecer a matéria indiferenciada, i.e., matéria não distinguível como diferentes substâncias tais como ferro, prata, ouro, etc. A física é a classe principal que trata com a matéria pura qua matéria. Na física podemos ter a necessidade de ter facetas-matéria baseadas em várias propriedades físicas como características. Há, por exemplo, a característica de estado-da-matéria. Sólidos, líquidos e gases são comumente conhecidos como facetas dentro daquela (RANGANATHAN, 1951, p.259).

(...) 372 Material

Na química, tecnologia e em várias das artes úteis, a categoria fundamental personalidade se manifesta como materiais tais como substâncias, produtos intermediários e produtos finais . (RANGANATHAN, 1951, p.263).

Outros autores, como Satija (2002), também observam esse aspecto do contexto influenciando a categoria a que pertence um conceito:

É sobre algo passivo que a faceta energia age. Por exemplo, na Biblioteconomia matéria é o tipo de documento, sejam livros, periódicos, manuscritos, microfilmes ou um CD. (...) Na indústria têxtil é o fio de tecido. Na música é o tipo de instrumento musical. Em economia é o papel ou o tipo de metal – o meio monetário. (SATIJA, 2002, p.28).

Dessa forma, pode-se considerar que os princípios adotados para a formação de assuntos, são também utilizados para nortear a formação de estruturas classificatórias. Kashyap (2001) corrobora essa afirmativa, ao destacar o fato de que os postulados de Ranganathan servem aos dois propósitos (além de colocarem o trabalho de classificação bibliográfica sob bases científicas):

O conjunto de postulados formulados por Ranganathan para a concepção de um sistema de classificação em bibliotecas, bem como para classificar os documentos bibliográficos de acordo com o seu assunto específico, coloca a teoria da classificação, e o trabalho de classificação e indexação dos documentos bibliográficos, em uma base científica sólida. (KASHIAP, 2001).

Nesse sentido, conclui-se que o contexto sempre se faz presente, e deve ser levado em consideração ao se elaborar uma estrutura classificatória e também para a formação de assuntos dos documentos, no momento da indexação.

2.2 O contexto na formação de assuntos

Ao se formar um assunto, as categorias de Ranganathan também devem ser sempre contextualizadas no domínio que se está analisando. Essa contextualização do domínio é sugerida por Ranganathan quando postula que cada assunto deve possuir um assunto básico como primeira faceta, e quando apresenta exemplos onde a categorização de um assunto varia de acordo com o contexto do assunto implícito na faceta básica. Dessa forma, é esse assunto básico que vai em um primeiro momento nortear a classificação dos conceitos pelas categorias. Por exemplo, se o assunto básico situa o domínio na área da Manufatura de Aço, então o conceito de ferro é considerado como pertencente à categoria de Matéria. Por outro lado, se o assunto básico situa o domínio na área da Metalurgia, então o ferro é considerado como pertencente à categoria de Personalidade, conforme exemplificado no QUADRO 1.

QUADRO 1 - Exemplos das categorias variando em contextos.

Entidade	como Personalidade em assuntos relativos a	Como Matéria em assuntos relativos a
Minério de ferro	Mineração	Metalurgia
Ferro	Metalurgia	Manufatura de aço
Aço	Manufatura de aço	Manufatura de chapa de aço
Chapa de aço	Manufatura de chapa de aço	Confecção de caixas de ferro
Caixa de aço	Confecção de caixas de ferro	Comércio de embalagens

Fonte: Traduzido/adaptado de: Ranganathan (1967a, p.407)

Essa ideia da vinculação das categorias ao contexto é apontada por vários autores (PIEIDADE, 1977; SILVA, 2010; ARTÊNCIO, 2007), porém deve-se observar que, pelo que podemos perceber, para Ranganathan o contexto parece influir na identificação da própria natureza do conceito. Por outro lado, como mencionado anteriormente, Ranganathan fornece exemplos de assuntos categorizados sem indicar o seu contexto, como é o caso de higiene, doença, etc. Esse fato, aliado com as definições por ele fornecidas para as suas categorias, onde o contexto não é mencionado, possivelmente pode ter dado margem a entendimentos equivocados de como classificar um assunto, levando a crer que a categoria pode ser pensada de forma independente de um contexto, pelo menos no que tange às categorias de Matéria e Personalidade.

Essa dependência do contexto para a categorização, como no exemplo do ferro (vide Quadro 1), sugere que o que se busca capturar é, de fato, o papel que este desempenha na formação de um assunto, dentro desse contexto: o ferro na Metalurgia tem um papel de elemento constituinte de outro. Entretanto, ao invés de pensar que existe uma relação entre coisas como, por exemplo, constituição, e que o ferro atua como um dos componentes dessa relação, Ranganathan opta pelo deslocamento da categorização do conceito: de Personalidade para Matéria. Ou seja, em vez de considerar que existem entidades distintas passíveis de categorização (o ferro e, por exemplo, um artefato constituído pelo ferro), Ranganathan assume que há apenas uma entidade, o ferro, que muda de natureza, de acordo com o contexto.

Essa é uma diferença importante que distancia as categorias de Ranganathan daquelas estudadas por Aristóteles na Filosofia, pois Aristóteles se propõe a estudar as coisas que existem no mundo dentro de uma perspectiva onde a categorização dessas coisas não está atrelada a um contexto que pode determinar sua mudança de natureza. Ou seja, no caso do ferro, este será pensado como o que de fato a sua natureza revela que ele é, ou seja, um elemento químico, independente de estarmos analisando o ferro no contexto da Manufatura de

ação ou da Metalurgia, embora esse metal se manifeste através de características diversas (que podem ser categorizadas), que vão fornecer, em conjunto, o seu conceito tal como o reconhecemos.

Dessa forma, o que se coloca é o entendimento do que é a coisa que está no mundo, embora inserida em um cenário em que são feitas afirmações sobre ela, utilizando-se de uma linguagem. Nesse sentido, observa-se a ideia de um elemento central, ou fundamental, aquele que exprime o que a coisa no mundo de fato é, e que é objeto da predicação:

O Ser se toma em múltiplos sentidos (...): num sentido, significa isto que a coisa é, a substância, e, em outro sentido, significa uma qualidade, uma quantidade ou um dos outros predicados deste tipo. Mas, entre todas estas acepções do Ser, é claro que o Ser em sentido primeiro é o “isto que é a coisa”, noção que não exprime nada além da própria substância.

(...) Também se poderia perguntar se o passear, o sentir-se bem, o estar sentado são ou não são seres; e da mesma forma em qualquer outro caso análogo: pois nenhum destes estados tem por si mesmo naturalmente uma existência própria, nem pode ser separado da substância, mas se há aí algum ser, será antes isto que passeia que é um ser, isto que está sentado, isto que se sente bem. (REZENDE, 2010, p.75).

Se, de fato, Ranganathan se inspirou em Aristóteles, não podemos nos furtar a considerar a *possibilidade* de certa analogia entre a ideia de substância, de Aristóteles, e a da categoria de Personalidade, de Ranganathan. Da mesma forma, poderíamos considerar como *possibilidade* que as outras categorias de Ranganathan seriam destinadas a comportar as possíveis manifestações de Personalidade, dela dependendo, o que, de fato, Ranganathan parece sugerir, quando menciona que as outras facetas de natureza diferente de Personalidade, são, na verdade, atributos de Personalidade (RANGANATHAN, 1951, p.263).

A questão da predicação, abordada por Aristóteles, também poderia *possivelmente* ser tomada por Ranganathan como inspiração para a identificação de uma sequência útil de ideias em um assunto. Como apoio para identificar a sequência útil de ideias em um assunto, Ranganathan utiliza o conceito de “Sintaxe Absoluta”, que propõe uma *sintaxe unificada* como alternativa para a variedade sintática que ocorre naturalmente nos diversos idiomas:

Por ‘Sintaxe Absoluta’ entende-se a sequência na qual as ideias das facetas de um assunto se arranjam nas mentes da maioria das pessoas. A Sintaxe Linguística é a Sintaxe das Palavras – isto é, a sequência na qual as palavras ficam dispostas em uma sentença ou no nome de um assunto em uma linguagem natural. A Sintaxe Linguística pode variar de acordo com a linguagem; geralmente varia (RANGANATHAN, 1967b, p. 38).

(...) Mas há apenas uma Sintaxe de Facetas para o assunto. Por essa

razão, conjectura-se que a Sintaxe de Facetas é a mesma que a Sintaxe Absoluta. Isso implica que a Sintaxe Absoluta é aquela que está de acordo com os Postulados e Princípios que guiam o projeto de uma Classificação Analítico-Sintética. (RANGANATHAN, 1967b, p. 40).

Essa influência da perspectiva dos assuntos é sugerida por Neelameghan (1975), contemporâneo de Ranganathan, quando descreve a categoria de Personalidade como “entidade ou objeto central de estudo” (*core entity of study or object of study*). De maneira análoga, Broughton (2006) também destaca que a categoria de Personalidade são os “conceitos centrais representando o objeto primário de estudo dentro de uma disciplina (...), uma vez que representavam a essência dessa disciplina”. Kashyap (2001), ainda, afirma que “parece que Ranganathan escolheu o termo ‘personalidade’ para representar o ponto focal da descrição ou objeto chave ou objetos de estudo ou descrição em uma obra”.

A noção da existência de um objeto central de estudo poderia justificar o porquê de um conceito como o ferro ser classificado de maneira diferente de acordo com o domínio. Nesse caso, sugerimos que embora o aspecto de ser uma propriedade ou ser uma matéria constituinte de algo seja importante, não é, por si só, um fator determinante para que se categorize um assunto na categoria de Matéria: é preciso que esse assunto não seja uma entidade central de estudo do domínio. Para isso, deve-se considerar qual é a ideia central, do domínio. Um exemplo que corrobora essa possibilidade é a classificação, por Ranganathan, de “propriedades da matéria” (que ele exemplifica como: sólido, vidro, cristal, líquido, gás), na Física, como pertencentes à categoria de Personalidade (RANGANATHAN, 1967b).

Neelameghan (1975) aponta os aspectos básicos da análise de assuntos em uma ilustração contendo os elementos da estrutura facetada de assuntos, onde também menciona o aspecto de elemento central de estudo (“entidade central de estudo ou objeto de estudo”), associado à categoria de Personalidade.

Com base nas questões apresentadas, procedemos a analisar alguns exemplos peculiares de classificação de assuntos fornecidos pelo próprio Ranganathan em algumas de suas obras que vão reforçar a ideia proposta da importância de se considerar a noção de um objeto central de estudo do domínio como elemento norteador da atividade de categorização de assuntos, no que concernem as teorias de Ranganathan.

3 ALGUMAS DIVERGÊNCIAS SOBRE AS CATEGORIAS DE RANGANATHAN

Conforme relatado previamente, é comum encontrar-se exemplos de conceitos classificados nas categorias de Ranganathan sem que o contexto seja definido de forma clara, incluindo exemplos dados pelo próprio Ranganathan. Entretanto, Ranganathan (1967a) deixa

clara a importância da questão do contexto, começando pela identificação da faceta básica, cuja identificação, segundo o autor, deve preceder o processo classificatório.

De forma coerente, a identificação do objeto central de estudo idealmente deveria ser feita a partir da definição do escopo de abrangência da faceta básica. No caso dos exemplos de Ranganathan, como não há definição, recorreremos à literatura e ainda, à própria identificação do autor para elementos pertencentes à categoria de Personalidade, partindo do pressuposto de que esta categoria expressa os objetos centrais de estudo da faceta básica, conforme apontado por Neelameghan (1975).

3.1 Alguns exemplos peculiares das categorias de Ranganathan

Os exemplos a seguir são fornecidos por Ranganathan no contexto da classificação de assuntos sejam eles explícitos ou implícitos, a partir da análise da obra. Conforme Ranganathan (1967a), os assuntos explícitos são os que se apresentam claramente a partir da análise do título do documento, enquanto que os implícitos dependem de uma análise preliminar do conteúdo do documento. Os assuntos implícitos serão destacados em itálico. Após o comentário sobre os assuntos, os conceitos neles identificados serão usados como exemplos para reflexão sobre questões que se colocam no pensar sobre o processo classificatório, onde a compreensão dos conceitos e suas naturezas ocupam posição central.

A estrutura da proteína e o microscópio eletrônico

Nesse assunto tem-se como faceta básica a Química. No caso desse exemplo, Ranganathan identificou as seguintes facetas isoladas, com suas respectivas categorias fundamentais: estrutura (matéria); proteína (personalidade); *determinação* (energia); microscópio eletrônico (matéria). Cabe lembrar que, de acordo com Ranganathan, a noção de faceta isolada é “alguma ideia ou complexo de ideias ajustadas para formar um componente de um assunto, mas que em si mesmo não é considerado um assunto” (RANGANATHAN, 1967a).

A classificação de microscópio eletrônico como matéria contrasta com a definição preliminar dessa categoria por Ranganathan, afinal não é nem material nem propriedade, nem, por si só, um método.

Entretanto, se considerarmos que o assunto básico é a Química, podemos perceber que o microscópio não se enquadra como um objeto central de estudo da Química. Mais especificamente, se consideramos o assunto apresentado acima, pode-se identificar que a proteína, enquanto uma molécula orgânica se constitui em um elemento focal de estudo. Aliado a isso, se tomarmos a definição mais recente da categoria Matéria, onde se inclui a noção de método (meio para atingir um fim), pode-se entender a classificação de Microscópio

como matéria, pois ele poderia ser um meio (embora enquanto instrumento) para analisar a estrutura da proteína. Um dos sentidos da noção de instrumento é justamente o de meio para obter algo, o que coaduna com a noção de método.

Diagnóstico por raio-x na criação de vacas

Nesse assunto tem-se como faceta básica a criação de animais domésticos. No caso desse exemplo, Ranganathan identificou as seguintes facetas isoladas, com suas respectivas categorias fundamentais: raio-x (matéria); diagnóstico (energia); doença (matéria); vaca (personalidade).

Raio-x, por si só, não é nem material nem propriedade, nem método. Diagnóstico por raio-x poderia ser um método, mas o conceito de raio-x tomado como isolado, não está de acordo com a noção usual de método, uma vez que raio-x é um tipo específico de radiação eletromagnética (SEARLE, 1990).

Entretanto, se considerarmos que o assunto básico é a criação de animais domésticos, pode-se perceber que o raio-x não se enquadra como um objeto central de estudo. Mais especificamente, se consideramos o assunto apresentado acima, podemos identificar que a vaca, enquanto um animal doméstico, é considerada um elemento focal de estudo. Nesse sentido, uma vez identificado que o raio-x não é objeto central de estudo, pode-se entendê-lo como um instrumento; no caso, um instrumento para apoiar a realização de um diagnóstico no animal, ou como um agente do processo de diagnóstico.

Registro em fita e proteção das canções folclóricas dos Todas

Nesse assunto tem-se como faceta básica a Sociologia. No caso desse exemplo, Ranganathan identificou as seguintes facetas isoladas, com suas respectivas categorias fundamentais: registro em fita (matéria); moribunda (matéria); proteção (energia); canções populares (matéria); Todas (personalidade).

Canção popular, por si só, não é nem material nem propriedade, nem método. Entretanto, se considerarmos que o assunto básico é a Sociologia, pode-se perceber que as canções populares não se enquadram como um objeto central de estudo. Isso considerando a conceituação de Sociologia como o “estudo objetivo das relações que se estabelecem (...) entre pessoas que vivem em uma comunidade ou num grupo social, ou entre grupos sociais diferentes que vivem no seio de uma sociedade mais ampla” (FERREIRA, 2009).

Mais especificamente, se consideramos o assunto apresentado acima, podemos identificar que os Todas, enquanto um grupo social (EMENEAU, 1988), tomam o lugar do elemento focal de estudo. Nesse sentido, uma vez identificado o objeto de estudo, pode-se

entender as canções possivelmente como sendo um instrumento para apoiar o relacionamento desse grupo, ou como um agente desse processo de relacionamento.

Controle de natalidade essencial na Índia hoje (1966)

Nesse assunto tem-se como faceta básica a Sociologia. No caso desse exemplo, Ranganathan identificou as seguintes facetas isoladas, com suas respectivas categorias fundamentais: *superpopulação* (matéria); *prevenção* (energia); controle de natalidade (matéria); Índia (espaço); hoje (tempo).

O conceito de superpopulação pode ser definido como “estado de uma região cujos recursos já não são suficientes para satisfazer as necessidades da população” (FERREIRA, 2009). Dessa forma, superpopulação, por si só, não é nem material nem método, mas pode ser uma propriedade, se concordarmos que estado é um tipo de propriedade.

Pode-se considerar um *estado* como um tipo de propriedade, e, nesse caso, pode-se considerar superpopulação como uma propriedade, mas há autores que estabelecem diferenças entre os conceitos de estado (aspecto dinâmico) e propriedade (aspecto estático). Kiwelekar e Joshi (2007) exemplificam as diferenças entre propriedade e estado. São propriedades de uma impressora: fabricante, localização, nome. São estados de uma impressora: ligada/desligada, ocupada/desocupada.

Se no caso do controle da natalidade considerarmos que o assunto básico é a Sociologia, pode-se perceber que o conceito de superpopulação não se enquadra como um objeto central de estudo. Cabe acrescentar que no exemplo do assunto em questão não é identificado nenhum conceito como pertencente à categoria de Personalidade. Sendo assim, a característica de superpopulação seria direcionada a um conceito central da faceta básica que estaria omissa no assunto (povo), mas que poderia ser identificado a partir da definição da faceta básica, no caso a Sociologia.

O exemplo do assunto acima é um caso singular, dentro do entendimento da categoria de Matéria, pois, dependendo de como se conceitua a noção de estado, não se constitui em nenhum dos aspectos considerados por Ranganathan (material, propriedade, método), nem mesmo na versão mais atual de sua Teoria da Classificação Facetada. Entretanto, como não se sabe, até onde se tem notícia, de qual exatamente foi o conceito de superpopulação adotado por Ranganathan, fica a dúvida de esse conceito ter sido tomado no sentido de uma qualidade ou característica e não de um estado, ou, então, de que Ranganathan considerava estado como sendo o mesmo que, ou um tipo de propriedade.

Controle de natalidade também, por si só, não é nem material nem propriedade. Se tomado de forma isolada, o conceito de controle poderia ser categorizado como Energia, pois

controle é o “ato de controlar”, portanto uma ação. Entretanto, poder-se-ia entender o controle de natalidade como um método (meio) para prevenir a superpopulação (apesar de que existe o conceito de métodos de controle de natalidade). Nesse caso, justifica-se a categorização de controle de natalidade na categoria de Matéria nos moldes propostos por Ranganathan, tomando-se a definição mais recente dessa categoria (RANGANATHAN, 1989). E, ainda, entende-se, nesse caso, a crítica de alguns autores à inclusão de método na categoria de Matéria, como mencionam Mazzocchi e Gnoli (2010), por conta de poder englobar conceitos de ação. Entretanto, se considerarmos o princípio de que a faceta básica é um contexto que influencia a determinação da classificação do conceito, então se pode considerar que, nesse caso, possivelmente o que ocorreu foi uma mudança de papel: de Energia para Matéria, ou seja, de ação para método. É um caso bastante peculiar, pois até então a dúvida se concentrava entre as categorias de Personalidade e Matéria e, a partir desse exemplo, percebe-se que a mudança de papéis atingiria pelo menos mais uma categoria, a de Energia. Entretanto, não se pode afirmar que houve, de fato, essa mudança de papéis, pois não se sabe qual foi o entendimento do conceito de controle de natalidade que Ranganathan adotou. Esse fato, corroborado por outras ocorrências de termos de entendimento dúbio, destaca a importância da definição de conceitos, como preconiza Dahlberg em sua Teoria do Conceito (1978).

4 SISTEMATIZAÇÃO DE QUESTÕES RELACIONADAS ÀS CATEGORIAS

Com base nas divergências discutidas na seção 3, propõe-se uma análise de questões que estão subjacentes às categorias de Ranganathan e sua aplicação. Para efetuar essa análise tomou-se como exemplo um dos assuntos abordados na seção 3.1, o qual contém conceitos das categorias de Personalidade, Matéria e Energia. Cabe destacar que a análise é feita a partir de um exemplo simples, sem a pretensão de ser exhaustiva, tendo como intuito introduzir questões que devem ser pensadas quando se propõe a criação de estruturas para a representação e organização de domínios do conhecimento. É importante ressaltar que o foco nesse caso não é a classificação bibliográfica e a identificação de assuntos, mas sim a identificação de conceitos (DAHLBERG, 1978) e seus relacionamentos. Nesse sentido, é natural que questões adicionais sejam introduzidas, pois o objetivo de uso das teorias de Ranganathan foi deslocado do seu propósito original.

4.1 Análise de conceitos relacionados às categorias de Personalidade, Matéria e Energia: a análise ontológica

O exemplo a ser discutido se refere ao assunto “estrutura da proteína e o microscópio eletrônico”. O objetivo que se deseja atingir é um entendimento mais preciso dos conceitos

envolvidos, sem a preocupação da identificação de assuntos para fins de classificação bibliográfica, mas sim para fins de representação do conhecimento.

Neste caso, propomos o que vem sendo denominada na literatura atual, como “análise ontológica”, ou seja, uma análise detalhada das propriedades dos objetos de um domínio visando à elaboração de modelos conceituais consistentes, onde atualmente as Ontologias de Fundamentação são as maiores representantes. (SMITH e KUMAR (2004); WELTY e GUARINO (2001)).

O microscópio eletrônico é um artefato tecnológico que pode ser usado (como instrumento) para visualizar a estrutura da proteína. Nesse caso, pode-se acrescentar que esteja desempenhando uma função (ARP e SMITH, 2008), pois ele foi construído de fato para observar coisas que não são observáveis (pelo menos não com facilidade) a olho nu. Entretanto, apesar de estar *desempenhando* uma função (observação de coisas microscópicas), o microscópio não é essa função, ele é um artefato.

Outro aspecto importante a destacar é que apesar da denominação da função denotar uma ação, a função não é uma ação em si, e sim o *potencial* de realizar essa ação, com um propósito específico. Dessa forma, adotou-se nomear a função do microscópio como “potencial de observação”, para não confundir com a ação de observação propriamente dita.

Mas será que existe uma dependência entre o artefato (enquanto portador) e a função? Ou seja, será que a função pode existir sem que o seu portador exista? E, ainda, será que o portador pode existir sem que a função exista?

Por outro lado, um microscópio eletrônico difere de um microscópio óptico comum em algumas características, uma delas sendo a utilização de feixes de elétrons em vez de luz para a observação do objeto de estudo (DEDAVID, MACHADO e GOMES, 2007), ou seja, a *técnica* utilizada para ampliação é diferente. Mas uma técnica é uma característica que não depende do microscópio. Ou seja, a técnica de feixe de elétrons não deixa de existir se o microscópio for destruído. O que é diferente do caso da função. Dessa forma, parecem existir características de natureza distinta, algumas dependentes e outras independentes de seu portador. Outra possibilidade seria considerar que o microscópio *seja* um método ou meio, situação na qual se poderia conceber que o microscópio, que naturalmente concebemos como um artefato, pode assumir outra *natureza* dependendo do contexto. Mas o que vem a ser a natureza de uma entidade? E, ainda, o que confere identidade a uma entidade? Natureza e identidade são coisas distintas? Por exemplo, é razoável pensar que um microscópio possua a natureza de Personalidade, mas a identidade de ser um artefato?

Uma estrutura pode ser considerada como uma propriedade de uma coisa, sendo essa coisa, no exemplo, a proteína. Ranganathan reconhece que a estrutura da proteína é algo de natureza distinta da proteína. Nesse caso, há que se pensar se a estrutura da proteína possui uma existência própria ou se ela depende da proteína e, ainda, se a proteína existe de forma independente de sua estrutura. De acordo com Alberts et al. (2002), a identidade de uma proteína está ligada à sua estrutura, e não pode ser entendida em separado dela (ALBERTS, et al. 2002, p. 125). Alberts et al. (2002) afirmam ainda que “a forma precisa de cada molécula de proteína determina a sua função em uma célula.”.

Nesse contexto algumas questões que se colocam são: (i) se as entidades podem possuir uma natureza mutável de acordo com o contexto; (ii) se existem entidades dependentes de outras entidades; (iii) se existem características que são essenciais a seus portadores; (iv) o que é natureza e identidade; (v) o que confere a identidade de uma entidade.

Uma determinação é uma ação e ações possuem aspectos temporais, ou seja, suas partes estão desdobradas no tempo. Isso contrasta com a noção de objetos concretos, cujas partes características estão presentes em todos os momentos no tempo e, ainda, com objetos de natureza abstrata, que independem de tempo e espaço, como números, por exemplo.

Dessa forma, entendemos que uma determinação é uma ação que possui partes (etapas) que se manifestam ao longo do tempo. Por exemplo, em um primeiro momento existe uma observação, depois uma análise, depois uma identificação de padrões, etc. A determinação é a soma de todas essas partes que ocorrem ao longo do tempo. Além disso, existe a questão da mudança ao longo do tempo. Podemos intuir a noção de objetos que mudam com o tempo, como, por exemplo, um microscópio que com o passar dos anos vai tendo sua lente desgastada (arranhada, embolorada). Por outro lado, um processo sendo o somatório de suas partes temporais, não comporta o conceito de mudança no mesmo sentido daquele dos objetos que persistem no tempo.

Outra questão, em relação à passagem do tempo, é que o tempo passa para os indivíduos, como, por exemplo, aquele microscópio específico do laboratório, com número de patrimônio 1234. Porém esse tempo não parece passar em relação ao conceito geral de microscópio, levando a crer que existem entidades individuais e entidades gerais e que o tempo as afeta de forma diferente.

Nesse contexto algumas questões que se colocam são: (i) se existem entidades que possuem partes no tempo e outras que não; (ii) se existem entidades que são afetadas de forma diferente ao longo do tempo; (iii) se existem entidades fora do tempo/espaço; (iv) se existem entidades gerais e individuais.

4.2 Resultados da sistematização

Espera-se que a sistematização das questões discutidas possa contribuir para o entendimento de aspectos que poderiam ser utilizados para uma caracterização mais precisa das categorias de Ranganathan e, eventualmente, considerar-se a adoção de outro conjunto de categorias, no contexto da elaboração de estruturas classificatórias para a representação e organização de domínios de conhecimento. Cabe destacar que nesse contexto tem sido comum a adoção de ontologias de fundamentação (CAMPOS, CAMPOS e MEDEIROS, 2011), que, de forma simplificada, pode-se considerar como sendo uma representação formal de um conjunto de categorias de alto nível e suas relações, o que se aproxima, de certa forma, da proposta de categorização de Ranganathan. Além disso, verificamos que para o entendimento das propostas de Ranganathan é necessário nos apoiamos em uma análise ontológica como apresentado na seção anterior.

De acordo com o que foi apresentado até então, diversas questões se colocam, podendo ser organizadas pelos seguintes aspectos: (i) esclarecimento de pressupostos; (ii) entendimento do que é o ser; (iii) entendimento do que é a natureza do ser. Espera-se, para cada aspecto, a compreensão de uma série de questões, cujas respectivas relevâncias são exemplificadas de forma sintetizada no QUADRO 2.

QUADRO 2: Sistematização de aspectos da análise dos conceitos da Seção 4.1

Aspecto / questão	Relevância
<i>(i) Pressupostos</i>	
Entender o que é natureza, identidade, tempo e espaço, deixando explícito o que se entende por cada uma dessas noções;	O entendimento ambíguo dessas noções pode gerar modelos que expressam ideias diferentes sendo tratadas como se fossem iguais, gerando um problema no que tange à sua interoperabilidade e compreensão (CAMPOS, 2005)
<i>(ii) Entendimento do que é o ser</i>	
Se as entidades podem possuir uma natureza mutável de acordo com o contexto;	Se uma entidade pode mudar de natureza, então é preciso conhecer quais os critérios que determinam essa mudança. Assumir que uma entidade pode mudar de natureza dificulta a interoperabilidade de modelos, pois os critérios envolvidos vão influir na determinação do que é o conceito.
O que confere a identidade de uma entidade;	A partir do conhecimento da identidade podem-se estabelecer relações mais precisas de equivalência, igualdade e subsunção entre conceitos. Pode-se ainda, definir a questão da constituição de uma coisa. Por exemplo, existem duas entidades no vaso de barro (o vaso em si e o barro), ou existe apenas o barro, que foi moldado no formato de um vaso? (CAMPOS, 2011).
Se existem características que são	Características essenciais são uma forma de determinar

essenciais a seus portadores;	a identidade de uma entidade (GUARINO, 1997).
<i>(ii) Entendimento do que é a natureza do ser</i>	
Se existem entidades gerais e individuais	Diferentes relações se aplicam a entidades gerais e individuais. Por exemplo: semelhança e amizade, entre indivíduos e gênero-espécie entre entidades gerais. Dessa forma, é importante saber o que se admite existir, assim como o que caracteriza uma e outra (BITTNER, DONNELLY, SMITH, 2004).
Se existem entidades dependentes de outras entidades	Assumir a dependência entre entidades implica, dentre outras coisas, o estabelecimento de relações entre elas (SMITH e GRENON, 2004), as quais, por coerência, devem estar representadas nos modelos do domínio.
Se existem entidades fora do tempo/espaço	Implica em se determinar de que maneira conceitos universais se situam em relação ao espaço e tempo (BRANQUINHO e SANTOS, 2013), ou seja, se estão todos fora de espaço-tempo ou se apenas alguns estão. E, ainda, em se considerar a existência de entidades abstratas, como, por exemplo, números (BITTNER e SMITH, 2003).
Se existem entidades que possuem partes no tempo e outras que não	Implica em determinar de que maneira se admite a mudança ao longo do tempo (GRENON e SMITH, 2004),

Pelo que pudemos perceber, para fins de representação do conhecimento existem aspectos que fazem com que seja importante entender e explicitar os pressupostos epistemológicos que estão fundamentando a identificação de cada categoria a ser utilizada.

É importante situar que epistemologia diz respeito à natureza do conhecimento, ou seja, a maneira como a realidade é percebida, medida e compreendida. Ela fornece bases teóricas em especial para as ontologias de fundamentação, que são baseadas em categorias de alto nível, e influenciam as escolhas que se encontram representadas nos conceitos nelas presentes (POLI e OBRST, 2009, p. 3). Para cada aspecto tratado pela epistemologia existem correntes ou teorias, que buscam explicá-los, sendo uma questão de escolha optar por uma ou outra. Essas escolhas, se explicitadas, fornecem condições para que se categorizem os conceitos de forma mais precisa, de acordo com os pressupostos que as sustentam.

5 CONCLUSÃO

As teorias de Ranganathan têm se mostrado úteis nos dias atuais, mesmo que hoje venham sendo utilizadas com o foco principal não na formação de assuntos, mas na representação e organização de conhecimento. Nesse contexto, muitas dúvidas têm sido levantadas, devido à dificuldade de se obter um consenso de como devem ser entendidas essas categorias. Até onde pudemos perceber, a classificação de conceitos nas categorias de

Ranganathan depende do contexto onde estão inseridos, deixando evidente que a natureza de um conceito pode variar o que é um elemento complicador para a correta identificação da sua categoria. Entretanto, dada a complexidade dos domínios que se apresentam como alvo dessa representação e da complexidade envolvida no trabalho intelectual de categorização, torna-se importante a necessidade de explicitar os pressupostos que são utilizados para definir cada categoria, de modo a minimizar o entendimento ambíguo.

A fim de abordar tal questão, apresentamos uma análise das divergências no uso das categorias de Ranganathan, bem como uma sistematização de questões suscitadas a partir dessa análise, mostrando a importância de se considerar o contexto na Teoria da Classificação Facetada e de se entender os pressupostos que norteiam a definição de categorias. Tanto na formação de assuntos, quanto na elaboração de estruturas classificatórias. Além disso, ressaltamos ainda que a metodologia utilizada para análise foi um exercício para demonstrar a possibilidade de uma análise ontológica sobre domínios e sua importância para a elaboração de estruturas classificatórias consistentes. Trabalhos futuros podem explorar uma proposta de definição mais precisa para as categorias de Ranganathan, de modo a deixar explícitos os pressupostos que as embasam.

REFERÊNCIAS

- ALBERTS, B.; JOHNSON, A.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WALTER, P. **Molecular Biology of the Cell**. 4th edition. New York: Garland Science; 2002.
- ARP, R.; SMITH, B. Function, role, and disposition in basic formal ontology. In: PROCEEDINGS OF BIO-ONTOLOGIES WORKSHOP (ISMB2008), p.45–48, 2008.
- ARTÊNCIO, L. M. **Princípios de categorização nas linguagens documentárias**. 2007. Dissertação de Mestrado. Escola de Comunicações e Artes, USP, São Paulo, 2007.
- BITTNER, T.; DONNELLY, M.; SMITH, B. Individuals, Universals, Collections: On the Foundational Relations of Ontology, In: PROCEEDINGS OF THE THIRD INTERNATIONAL CONFERENCE, Amsterdam: IOS Press, p. 37–48, 2004.
- BITTNER, T.; SMITH, B. Directly Depicting Granular Ontologies, In: PROCEEDINGS OF THE 1st INTERNATIONAL WORKSHOP ON ADAPTATIVE MULTIMEDIA RETRIEVAL, Hamburg, 2003.
- BRANQUINHO, J.; SANTOS, R., orgs. - Regressões ao InfInItto em metafísica. **Compêndio em Linha de Problemas de Filosofia Analítica**. ISBN: 978989855322-5, 2013.
- BROUGHTON, V. The need for a faceted classification as the basis of all methods of information retrieval. **Aslib Proceedings**, v. 58, n.1/2, p. 49-72, 2006.
- CAMPOS, L. M. **Diretrizes para definição de recorte de domínio no reuso de ontologias biomédicas: uma abordagem interdisciplinar baseada na análise do compromisso ontológico**. 2011. Tese. UFF/IBICT, Rio de Janeiro, 2011.

CAMPOS, M. L. A. ; CAMPOS, L. M. ; MEDEIROS, J. . A Representação de Domínios de Conhecimento e uma Teoria de Representação: a ontologia de fundamentação. **Informação & informação (UEL. Online)**, v. 16, p. 140-164, 2011.

CAMPOS, M. L. A. **A problemática da compatibilização terminológica e a integração de ontologias**: o papel das definições conceituais. In: VI ENANCIB. Encontro Nacional de Pesquisa em Ciência da Informação, Florianópolis, 2005.

CAMPOS, M. L. A.; GOMES, H. E. Taxonomia e classificação: o princípio de categorização. **DataGramZero**, v. 9, n. 4, ago. 2008.

CAMPOS, M.L.A. Linguagem Documentária: teorias que fundamentam sua elaboração. Niterói, EdUFF, 133p, 2001.

CAMPOS, M.L.A.; GOMES, H.E.; OLIVEIRA, L.L. As Categorias de Ranganathan na organização dos conteúdos de um portal científico. **DataGramZero**, Rio de Janeiro, v. 14, n. 3, jun. 2013.

DAHLBERG, I. A referent-oriented analytical concept theory of interconcept. **International Classification**, v. 5, n. 3, p. 142-150, 1978.

DEDAVID, B. A.; MACHADO, G. GOMES, C.I. Microscopia Eletrônica de Varredura: aplicações e preparação de amostras, EDIPUCRS, 2007.

EMENEAU, M.B. A Century of Toda Studies. **Journal of the American Oriental Society**. v. 108, n. 4, p. 605-609, 1988.

FERREIRA, A.B.H. **Novo dicionário língua portuguesa**, 4a. Ed. Curitiba. Positivo, 2009.

GISS, M. D.; WILD, P. J.; McMAHON, C. A. The generation of faceted classification for use in the organization of engineering design documents. **International Journal of Information Management**. v. 28 n. 5, p. 379-390, 2008.

GRENON, P.; SMITH, B. SNAP and SPAN: Towards Dynamic Spatial Ontology. **Spatial Cognition and Computation**, v.4, n.1, p. 69-103, 2004.

GUARINO, N. Some Organizing Principles for a unified top-level ontology. In: **Spring Series on Ontological Engineering**. Stanford, p.57-63, 1997.

KASHYAP, M. Similarity between Ranganathan's postulates and Chen's entity-relationship approach to data modeling and analysis. **DESIDOC Bulletin of Information Technology**, v.21, n.3, p.3-16, 2001.

KIWELEKAR, A.W., JOSHI, R.K.: An object oriented metamodel for bunge-wand- weber ontology. In: In PROCEEDINGS OF SWeCKa 2007, Workshop on Semantic Web for Collaborative Knowledge Acquisition at IJCAI, 2007.

LA BARRE, K. Facets, search and discovery in next generation catalogs: Informing the future by revisiting past understanding. In: PROCEEDINGS OF THE 11th INTERNATIONAL CONFERENCE OF THE ISKO, p. 269-275, 2010.

LOUX, M.J. **Metaphysics a contemporary introduction**, 3rd. edition, New York, NY, 2006.

MARKIE, P. **Rationalism vs. empiricism**. In Zalta E.N. (Ed.), The Stanford encyclopedia of philosophy, 2004.

MAZZOCCHI, F. GNOLI, C. S.R. Ranganathan's PMEST Categories: Analyzing their Philosophical Background Cognitive Function. **Information Studies**, v.16, p. 133-147, 2010.

NEELAMEGHAN, A. Absolute syntax and structure of an indexing and switching language. In: PROCEEDINGS OF THE 3rd. INTERNATIONAL STUDY CONFERENCE ON CLASSIFICATION RESEARCH, p.165-176. Bangalore, 1975.

PIEIDADE, M. A. Requião. **Introdução à teoria da classificação**. Rio de Janeiro: Interciência, 1977

POLI, R.; OBRST, L. The Interplay Between Ontology as Categorical Analysis and Ontology as Technology. in R. Poli, M. Healy, A. Kameas, eds., **TAO-Theory and Applications of Ontology**, Springer 2009.

RANGANATHAN, S.R. **Colon Classification**, 7th. Edition, edited by M.A. Gopinath. Bangalore: Sarada Ranganathan Endowment for Library Science, 1989.

RANGANATHAN, S.R. Hidden roots of classification. **Information Storage and Retrieval**, v.3, n.4, p. 399-410, 1967b.

RANGANATHAN, S.R. **Prolegomena to Library Classification**. Bombay: Asia Publishing House, 1967a.

RANGANATHAN, S.R. **Classification and Communication**. Sarada Ranganathan Endowment for Library Science. University of Delhi, 1951.

REZENDE, A. (Org.). Curso de Filosofia: para professores e alunos dos cursos de segundo grau e de graduação. Rio de Janeiro: Zahar, 15ª. Edição, 2010.

SATIJA, M.P. **Manual of Practical Colon Classification**, 4th Ed., New Delhi, India, 2002.

SEARLE, J R. Is the brain's mind a computer program? **Scientific American**, v. 262, n.1, p. 26-31, 1990.

SILVA, A.R. **Estudo dos princípios de categorização na Biblioteconomia e Ciência da Informação**: Ranganathan - entre a teoria clássica e a abordagem cognitiva contemporânea. 2010. Dissertação de Mestrado. Universidade Federal de Minas Gerais, Belo Horizonte. 2010.

SMITH, B.; GRENON, P. The Cornucopia of Formal-Ontological Relations, **Dialectica**, v. 58, n.3, p. 279-296, 2004.

TENNIS, J.T. Ranganathan's layers of classification theory and the FASDA model of classification. In: PROCEEDINGS FROM NORTH AMERICAN SYMPOSIUM ON KNOWLEDGE ORGANIZATION, v. 3. Toronto, Canada, p. 185-195, 2011.

XV Encontro Nacional de Pesquisa em Ciência da Informação – **ENANCIB 2014**

MAPEAMENTO TEMÁTICO: O GT2 DO ENANCIB

THEMATIC MAPPING: GT2 OF THE ENANCIB

Rosana Portugal Tavares de Moraes
 Maria Luiza de Almeida Campos

Resumo: A atividade de acompanhamento das construções discursivas de um determinado domínio do saber é essencial para o trabalho de Organização do Conhecimento. Conhecer o modo como um domínio se estrutura conceitualmente possibilita um olhar avaliativo sobre o tratamento dado às suas temáticas de pesquisa e se mostra útil para que seus membros conheçam e avaliem seu fazer científico. Esta comunicação demonstra o estado da arte do domínio da “Organização e Representação do Conhecimento”, por meio da análise das palavras-chave das comunicações orais do Grupo de Trabalho 2, do Encontro Nacional de Pesquisa em Ciência da Informação, no período compreendido entre 2003-2013. Após a observação de alguns critérios necessários ao manuseio dos dados, os termos foram classificados no “Sistema de Classificação para Literatura de Organização do Conhecimento”. Os dados foram representados através de mapas com o delineamento conceitual do domínio, efetuando-se em seguida uma comparação com dados de outras pesquisas que utilizaram este mesmo modelo. A pesquisa demonstrou a preferência pela pesquisa aplicada e a carência de pesquisas referentes aos aspectos metodológicos utilizados no domínio, bem como à Representação Descritiva da Informação.

Palavras-chave: Mapeamento temático. Análise de Domínio. Organização e Representação do Conhecimento. GT2. ENANCIB.

Abstract: The activity of following the course of discursive constructions of a determined knowledge domain is essential to the task of Organization of Knowledge. Knowing how a domain is conceptually structured makes an evaluative view possible on how its research themes are dealt with and is useful for allowing fellow members to become familiar with and evaluate their scientific métier. This communication shows the state of the art of the “Organization and Representation of Knowledge”, by way of analyzing keywords of oral communications presented in Workgroup 2 of the National Meeting on Research in Information Science, from 2003-2013. After observing certain criteria necessary for manipulating data, the terms were classified in the “System of Classification for Literature on the Organization of Knowledge”. Maps were used to represent the conceptual delineation of a domain, followed by a comparison of data from other research that also employed this model. The research demonstrated a preference for applied research and a lack of research making reference to the methodological aspects used in the domain, and also the Descriptive Representation of Information.

Keywords: Thematic mapping. Domain Analysis. Organization and Representation of Knowledge. GT2. ENANCIB.

1 INTRODUÇÃO

Conhecer os temas de pesquisa dos domínios do conhecimento constitui uma atividade essencial ao trabalho de Organização do Conhecimento. Cada domínio tem um conjunto de conceitos utilizados em seus processos comunicativos que o distingue dos demais. Alguns são próprios, exclusivos, outros são utilizados por diversos domínios. Contudo, em cada um deles, podem ter um sentido singular. A identificação deste sentido específico dos conceitos é um

dos objetivos do trabalho de Organização do Conhecimento para uma recuperação eficaz da informação. Sendo assim, trata-se de uma prerrogativa do próprio domínio sistematizar os meios de se extrair e conhecer a semântica dos termos envolvidos nos objetos, nos processos e nos métodos utilizados na produção discursiva dos demais domínios do conhecimento.

Em tempos de efervescência de fusão e cisão de disciplinas, com formação de novos domínios do saber e novas configurações discursivas, é importante que a Ciência da Informação acompanhe os caminhos da formação dos discursos científicos, aproveitando-se da meta-informação oriunda das atividades de Organização e Representação da Informação e do Conhecimento.

Essas coordenadas informacionais acerca do conhecimento, não apenas seriam úteis para otimizar a comunicação científica, como instrumentos para a recuperação e disseminação da informação, mas também se mostrariam significativas para o monitoramento e análise da produtividade científica de organizações, instituições de pesquisa, regiões e países, numa re-utilização dessa meta-informação para a gestão político-administrativa das atividades científicas. [...] Pode dizer-se, aliás, que a ciência da informação tem sido uma das guardiãs da comensurabilidade dos discursos científicos, caracterizando a ciência como a prática social de maior potência de circulação informacional, da maior eficácia de codificação e explicitação de seus próprios processos produtivos e de maior rigor e eficácia institucional na regulamentação dos mecanismos e práticas de estabilização de discursos e saberes (GONZÁLEZ DE GÓMEZ, 2001, p. 14, grifo nosso).

No domínio da Organização e Representação do Conhecimento é possível notar a coexistência de diversas correntes conceituais. Essa afirmação encontra respaldo na pesquisa da professora Mariângela Fujita (2008), que analisou as comunicações apresentadas no Grupo de Trabalho (GT2) do Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB), no período compreendido entre 2005 e 2007. A este respeito, ela declara:

[...] a área está em pleno vigor, dado suas diversidades conceituais teóricas e metodológicas em discussão. Portanto, se por um lado, falta à área de Organização do Conhecimento uma consolidação científica, de outro, é inegável sua diversidade conceitual e o impacto de seus resultados para a Organização do Conhecimento de outras áreas científicas. Nesse contexto, é absolutamente fundamental à comunidade científica conhecer seu domínio em Organização do Conhecimento [...] (grifo nosso)

Partindo da mesma perspectiva da pluralidade conceitual no domínio, mas tomando como base os trabalhos apresentados nos encontros internacionais e regionais da International Society for Knowledge Organization (ISKO), Gomes (2009, p. 61) ressalta que: “[...] não ficam claras as fronteiras da OC, o que é uma característica da atividade científica atual, ou seja, não há limites claros nas atuais atividades científicas da área. Mas haveria um núcleo que caracterizaria a área”. A autora deixa clara a existência de um corpo basilar do domínio, a

dinâmica conceitual entre o núcleo e as bordas consiste em um movimento contínuo impulsionado pelos fenômenos pesquisados.

Com base nas comunicações orais do GT2 “Organização e Representação do Conhecimento” do Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB), esta pesquisa visa à identificação dos assuntos tratados neste evento, mais precisamente em suas dez últimas edições, período compreendido entre 2003 e 2013. Por meio do discurso produzido, pretende-se investigar a configuração temática das comunicações orais apresentadas, ordenando-as em um modelo classificatório. Neste caso, o modelo escolhido foi o “Sistema de Classificação da Literatura de Organização do Conhecimento”, especializado na área de Organização do Conhecimento.

O estudo de um domínio do conhecimento, ao identificar e classificar os seus conceitos, de forma a visualizar sua distribuição em uma tabela, pressupõe um olhar avaliativo sobre o seu discurso, ainda que delimitado por um esquema de classificação. A presente pesquisa propõe uma reflexão sobre o próprio domínio, reflexão esta que permitiria identificar o *corpus* de assuntos abordados, relacionados aos seus objetos/questões de pesquisa.

Esse “olhar para dentro” é essencial para qualquer domínio científico que, por meio da análise e avaliação do seu discurso, pode corrigir deficiências e avançar em pesquisas de qualidade. Marteleto (2009, p. 20) baseou-se em Bourdieu³⁹ (2001), que denominou essa atividade “reflexividade” para se referir à “[...] necessidade de submeter a ciência a uma análise histórica e sociológica, que permite àqueles que a fazem compreender os mecanismos sociais que orientam a prática científica e a própria estrutura e dinâmica do seu campo disciplinar.”

Esta comunicação consiste no resultado parcial da pesquisa de mestrado desenvolvida no PPGCI/UFF e demonstrou outros aspectos além do aqui apresentado. No entanto, para esta comunicação priorizou-se demonstrar o estado da arte das comunicações orais do GT2 representadas pelas palavras-chave atribuídas por seus autores.

É preciso reconhecer que tais palavras-chave nem sempre representam fielmente o conteúdo do documento. Contudo, julgou-se que esta seria a melhor forma de se analisar as temáticas de um determinado domínio. Pois, uma vez que a indexação constituiria uma tarefa árdua, impossível de ser feita manualmente, e estaria sob o crivo ou o julgamento do

³⁹ BOURDIEU, Pierre. **Science de la Science et réflexivité**. Cours du Collège de France 2000-2001. Paris: Raisons d’Agir, 2001.

indexador. A indexação automática demandaria o estudo de softwares específicos e, também, não constitui um método isento de críticas. Portanto, diante da análise destas opções, escolheu-se trabalhar com as palavras-chave, tal como atribuídas pelos autores das comunicações. Esta opção deveu-se tanto por entendermos que o autor é a melhor pessoa para atribuir termos representativos a seus escritos, como por estes últimos apresentarem maior consistência para o tratamento e a interpretação dos dados, ainda que esta fosse uma “escolha de Sofia”.

O caminho teórico escolhido como base de apoio para pesquisa foi a construção teórica da própria área. Aqui são apresentadas algumas considerações sobre as questões que envolvem as unidades de representação em um dado domínio, como também procedimentos de análise para o recorte conceitual de um domínio do conhecimento e aspectos da metodologia de Análise de Domínio. Os termos classificados no “Sistema de Classificação para Literatura de Organização do Conhecimento” permitiram visualizar por meio de mapas de cada classe, o delineamento conceitual do domínio de Organização e Representação do Conhecimento no GT2. Além disso, foi possível comparar estes dados com os de outras pesquisas que utilizaram o mesmo modelo conceitual, mas com outros tipos documentais. Tal comparação permitiu demonstrar e avaliar a organização conceitual do domínio do GT2 diante de outras fontes de pesquisas na área de Organização do Conhecimento.

2 UNIDADES DE REPRESENTAÇÃO DOS ASSUNTOS DE DOCUMENTOS

A expansão e a crescente utilização da informação em meio digital maximizaram a importância das palavras-chave como indicativas de conteúdo informacional. Com diferentes propósitos, elas são determinantes no processo de refinamento de busca e na recuperação de dados em sistemas de informação.

A atribuição de palavras-chave tem o objetivo de representar o conteúdo dos documentos em um duplo esforço: primeiro, o de representar fielmente os assuntos abordados no texto e, segundo, o de buscar uma aproximação semântica com o vocabulário utilizado pelo usuário como termo de busca.

A palavra-chave pode ser retirada do texto, como também de um vocabulário controlado, podendo demonstrar os aspectos gerais ou específicos do texto. Em ambos os casos, estará mais próxima da linguagem do especialista, ou seja, da terminologia utilizada em sua área de assunto. Esta é a constatação de Borba, Van der Laan e Ros Chini (2012, p. 29).

A palavra-chave passa a ser considerada um termo ou descritor, exatamente quando sofre um controle terminológico, no qual o conteúdo é representado pelo termo exato, consensualmente autorizado pela área de assunto.

Ao comentarem a complexidade de se representar o conhecimento, Gopinath e Das (1997) definem representação do conhecimento como “[...] um método para expressar o pensar humano e as concepções que os homens vêm reunindo através das diferentes épocas e regiões.” Esse conhecimento, oriundo do campo das ideias e do pensar, ao ser registrado, encaixa-se no tipo de conhecimento objetivo, definido por Gopinath e Das (1997) como: “conhecimento registrado nos documentos e apresentado de maneira lógica e sistemática para rápida absorção intelectual.”

Dahlberg (1993, p. 211) enuncia que: “o conhecimento por si só não pode ser apreendido ou representado a menos que seja apresentado por unidades do conhecimento e suas possíveis combinações em palavras/termos ou afirmações.” Dessa forma, a autora ressalta que qualquer tarefa de organizar o conhecimento deve ter como base as unidades do conhecimento, que nada mais são do que os conceitos, definidos como: “Uma unidade de conhecimento (conceito) é a síntese das características essenciais de um referente a ser representado por designações (termos, nomes, códigos)” (DAHLBERG, 2011, p. 69). Dahlberg ainda adverte que “conceito” tem sido entendido como “o significado de uma palavra” e tratado dessa forma na literatura da área. Este é seu sentido linguístico, contudo, para os interesses da Organização do Conhecimento, importa a concepção de representação de uma dada realidade demonstrada pelo conceito (Id., 2011, p. 69).

3 O RECORTE CONCEITUAL DE UM DOMÍNIO

Quando se propõe a olhar um domínio do conhecimento, com o objetivo de construir uma estrutura conceitual, o profissional da informação lança-se em uma atividade de modelagem do conhecimento, uma tarefa extremamente complexa. Souza (2012, p. 4) percebe os diferentes olhares do universo do conhecimento como decorrentes das funções práticas que as classificações exercem. Diferentemente dos pesquisadores e administradores da ciência, os profissionais da informação, por sua vez, estão preocupados com a organização e a representação com o objetivo de buscar e acessar o conhecimento registrado. Com diferentes propósitos, cada segmento de usuários lança um olhar específico sobre o conhecimento que deseja entender, organizar, classificar. Com esse entendimento, Souza (2012, p. 5) declara que: “[...] podem existir tantas classificações do universo do conhecimento quantos possam ser os objetivos de organizar e representar esse conhecimento.”

Toda essa tarefa é antes de tudo uma construção mental, na qual o objeto físico (como por exemplo: a classificação de plantas ou de animais) ou mental (conceitos, teorias, etc.), alvo da representação, é pensado em todo o seu significado, a sua relação com outros objetos, o modo como é interpretado, onde costuma ser aplicado e, para os interesses do profissional da informação, qual a melhor forma de encaixá-lo em um Sistema de Organização do Conhecimento, com vistas à representação e sua posterior recuperação.

Esse processo de representação de uma dada realidade pode ser entendido como um modelo, que é definido por Sayão (2001, p. 83) como “[...] representações simplificadas e inteligíveis do mundo, que permitem vislumbrar características essenciais de um domínio ou campo de estudo.” O autor ressalta ainda que os modelos consistem em uma compreensão aproximada dos fenômenos e, como tal, devem ser considerados como uma representação, jamais constituindo a realidade em si, pois como: “[...] aproximações altamente subjetivas, no sentido de não incluírem todas as observações e mensurações e medições associadas, mas, como tais, são valiosas por ocultarem detalhes secundários e permitirem o aparecimento dos aspectos fundamentais da realidade.”

Diante dessa complexidade, cabe ao profissional cercar-se de todas as informações possíveis sobre o domínio que deseja classificar. Também deve avaliar se é importante seguir os modelos classificatórios existentes, se cabe uma adaptação à suas necessidades, ou se deve construir estruturas classificatórias específicas, de modo a melhor representar o conhecimento de um determinado domínio. Nesse sentido, Chernyi (1997, p. 6) assegura que: “O propósito é o fator determinante para a escolha de um método de descrição, formalização e representação do conhecimento.”

Assim, a representação de um domínio do conhecimento implica um recorte conceitual, quando intencional e propositadamente seleciona aspectos, funções, motivos, resultados estruturados em termos ou conceitos, os quais devem ser organizados em uma sequência lógica de entendimento, em comum acordo com o propósito ou com o principal objetivo da representação.

4 PROCEDIMENTOS DE ANÁLISE DE DOMÍNIO

No campo da Ciência da Informação os estudos sobre Análise de Domínio foram sistematizados por Birger Hjørland e Hanne Albrechtsen, em 1995, como um paradigma na área. Neste sentido, os autores afirmam que: “[...] a melhor forma de entender a informação é através do estudo de domínios do conhecimento através das comunidades discursivas, que são partes da divisão de trabalho na sociedade” (1995, p. 400, tradução nossa).

Hjørland (2004, p. 18, tradução nossa) destaca que a Análise de Domínio “[...] não está preocupada com o usuário em geral, vendo-o como pertencente a diferentes culturas, estruturas sociais e domínios do conhecimento.” O autor relata ainda que pessoas envolvidas com a atividade de informação, tais como produtores, usuários ou intermediários, estão ligadas em maior ou menor grau a comunidades, que compartilham uma linguagem em comum, gêneros e outras práticas informacionais típicas.

Trata-se de uma proposta metodológica que busca entender a informação no ambiente de sua criação e uso, por meio das pessoas envolvidas e de suas formas de comunicação em espaços específicos. Estes ambientes, com suas regras, formalizações, padrões e convenções, dentre outros, formam uma comunidade discursiva que, de acordo com a atividade exercida, requer maior e mais complexo grau de informação. Para Hjørland e Pedersen (2005a, p. 590) “[...] estudos de diferentes domínios, tradições, gêneros, e perspectivas na literatura, podem auxiliar a identificação de padrões que podem ser úteis na recuperação da informação.”

A ênfase da Análise de Domínio recai sobre o tratamento do assunto do documento visando à sua recuperação. Desta perspectiva, é preciso repensar o fazer do profissional indexador, se antes tinha sua atenção voltada para a linguagem apresentada no item, alvo da indexação, e para a melhor representação possível - através dos termos de tesouros e listas de cabeçalhos de assunto - agora também precisa conferir atenção especial à linguagem utilizada pela comunidade de falantes na área temática do item que está sendo indexado, além dos possíveis usos sociais ou aplicações práticas que a comunidade fará do mesmo.

Hjørland (2005b) defende que a melhor forma de se conhecer um domínio é identificar as teorias ontológica, epistemológica e sociológica em voga. A presente pesquisa pautou-se na abordagem ontológica, que envolve explicações dos seguintes termos: áreas, domínios, elementos, campos, tipos, objetos, problemas, assunto (substância/diferença substancial) e tópicos. As relações genéricas do tesouro e os sistemas de classificação constituem exemplos de relação ontológica (HJORLAND, 2005b, p. 132).

Com o objetivo de se conhecer um domínio, Hjørland (2002a) aponta onze abordagens - que podem ser entendidas como orientações - que direcionam o profissional da informação em relação a quais documentos consultar, em sua busca por identificar e delimitar um domínio que deseja conhecer. Nesta pesquisa será utilizada a segunda abordagem, que trata da classificação especializada e visualiza de que modo os conceitos e as relações semânticas estão configurados através de um modelo conceitual determinado. Apesar de Hjørland (2002a) apresentar estas abordagens, notou-se que se restringem aos tipos documentais que deverão ser consultados, portanto se faz necessário especificar as ações metodológicas de

como analisar os termos, de forma a extrair e explicitar o máximo de conhecimento representado no domínio. Sendo assim, buscaram-se os princípios metodológicos da análise conceitual apoiado em Dahlberg (1978, 1979).

Para o entendimento de um domínio, é essencial o estudo da comunidade discursiva, de como se configura seu processo de comunicação, sob quais critérios se constitui, como se desenvolve e quais fatores a influenciam ou chegam a alterá-la. O reconhecimento de uma comunidade discursiva se dá por quem está fora dela, sua legitimidade sendo conferida pela percepção do discurso produzido e pela forma como interpreta e utiliza as teorias e os conceitos.

5 CONTEXTUALIZAÇÃO DO MARCO EMPÍRICO E PROCEDIMENTOS METODOLÓGICOS

Foram utilizadas as palavras-chave das comunicações orais do GT2 do ENANCIB, no período de 2003 a 2013, mais precisamente, das dez últimas edições do evento. O ENANCIB é o principal evento científico da área da Ciência da Informação, configurando-se como um espaço privilegiado de discussões acadêmicas, no qual os alunos dos Programas de Pós-Graduação do país têm a oportunidade de apresentar o resultado de suas pesquisas de Mestrado e Doutorado e os pesquisadores da área de relatar seus estudos. Por este motivo, trata-se de um evento que possibilita um compartilhamento singular de conhecimentos e aprendizagem.

Com o intuito de conferir maior consistência à interpretação dos resultados, foi necessária a adoção de alguns critérios. Sendo assim, foram consideradas as comunicações orais de origem nacional e os autores que apresentaram mais de três trabalhos no período compreendido pela pesquisa. Logo, dos 229 autores selecionados para o período chegou-se a 43 autores. Pode-se afirmar que estes últimos constituem a comunidade discursiva do GT2, sendo responsáveis por 167 comunicações orais com 734 palavras-chave, entre as quais verificou-se a co-ocorrência de similaridade lexical, chegando-se a um total de 438. Neste conjunto, foram excluídas 121 palavras-chave consideradas vazias de sentido, as restantes sendo submetidas a uma análise conceitual para identificação de palavras compostas, sinônimas e homônimas. Sendo assim, chegou-se a um total de 290 termos, que formaram o corpus de análise da pesquisa.

Para classificação destes termos no Sistema de Classificação para Literatura de Organização do Conhecimento foi necessário entender cada um deles. Sendo assim, confeccionou-se uma Ficha Terminológica para o registro dos principais dados, tais como:

título da comunicação oral, explicação/definição de cada termo, dados de filiação institucional e ano do ENANCIB. Para o seu preenchimento, priorizou-se a explicação do termo apresentada pelo autor. Quando não identificada tal explicação, foram consultados Dicionários, Glossários e a Biblioteca Digital de Teses e Dissertações.

Os termos foram classificados no Sistema de Classificação para Literatura de Organização do Conhecimento, modelo classificatório idealizado por Ingtraut Dahlberg, trata-se do primeiro sistema de classificação especializado, criado especialmente para o domínio, que vem sendo utilizado ininterruptamente desde 1974 para classificar os artigos deste periódico, os trabalhos apresentados na ISKO, bem como aqueles do *Boletim Bibliográfico*, publicação que lista as referências de documentos na área de Organização do Conhecimento. Este sistema tem como objetivo apresentar o escopo das temáticas desenvolvidas na área, em um esforço de representar os assuntos estudados e demonstrar os possíveis desdobramentos de cada um deles. Dahlberg (1993, p. 211) afirma que a divisão da tabela compreende três grandes eixos temáticos: o da Teoria dos Conceitos (classes 1 a 3); o da Terminologia Sistemática (classes 4 a 6) e o da Organização do Conhecimento (classes 7 a 9). A classe 0, específica para a subdivisão de forma, não foi utilizada por não atender aos propósitos desta pesquisa.

6 MAPEAMENTO TEMÁTICO DAS COMUNICAÇÕES ORAIS APRESENTADAS NO GT2

Os termos classificados podem ser visualizados por meio de um mapa com cada classe da tabela (Figuras 1 a 8). Na parte central do mapa está o título da classe, ligado por uma linha contínua de cor vermelha aos títulos da subclasse, ligados por sua vez aos termos. Para facilitar a visualização e o entendimento do modo como os termos foram distribuídos em cada classe, somente foram demonstradas no mapa a primeira e a segunda subdivisão da tabela com letra maiúscula. Os termos estão destacados em diagrama na cor correspondente à classe à qual pertencem - os que se referem à pesquisa aplicada estão apresentados em fonte branca, diferente dos demais, em fonte preta.

Na FIGURA 1 está representada a primeira classe – Fundamentos teóricos e problemas gerais. Nela, foram observadas 107 frequências nos 42 termos analisados. A análise desses dados e a visualização do mapa permitem reconhecer maior diversidade de temáticas na subclasse 12 – “Conceito e Organização do Conhecimento”, como também maior frequência (37 vezes). Este fato demonstra a preocupação da área com as questões que envolvem o conceito e seu tratamento na Organização do Conhecimento.

FIGURA 1: Classe 1 – Fundamentos teóricos e problemas gerais.



Fonte: Dados da pesquisa.

Na FIGURA 2, está representada a segunda classe, intitulada “Estrutura e construção de Sistema de Classificação e Tesouro”, que apresentou 38 termos com 110 frequências. Nela, sobressai a subclasse 21, “Questões gerais de Sistema de Classificação e Tesouro. Linguagens de Informação em geral”, com maior número de frequência (72 vezes) e maior especificidade (15 termos). “Ontologia”, um dos termos com maior frequência de estudo na área é representado nesta subclasse, na qual também se observa o detalhamento das temáticas envolvidas no assunto.

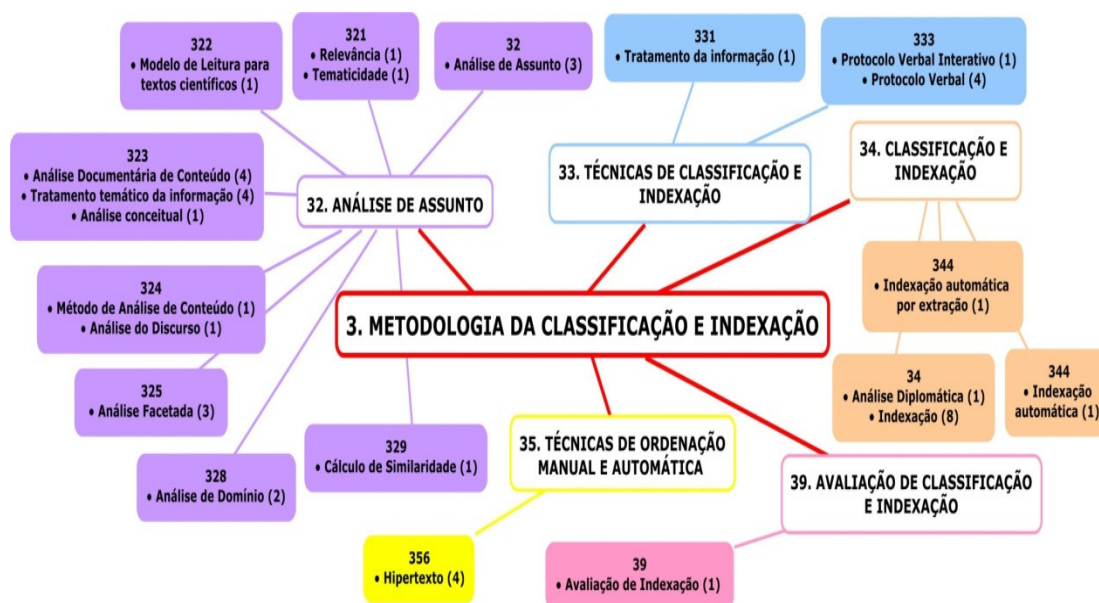
FIGURA 2: Classe 2 – Estrutura e construção do Sistema de Classificação e Tesouro



Fonte: Dados da pesquisa.

A terceira classe está representada na figura 3. Nela, são apresentados 21 termos com 48 frequências. Esta classe trata dos aspectos metodológicos envolvidos no trabalho de classificação e indexação. Nota-se que a subclasse 32 “Análise de assunto” reúne o maior quantitativo de pesquisa: foram 12 termos com 23 de frequência. Este dado demonstra a preocupação do domínio com métodos que viabilizem a identificação do assunto dos documentos, além da identificação das principais metodologias utilizadas nas pesquisas da área.

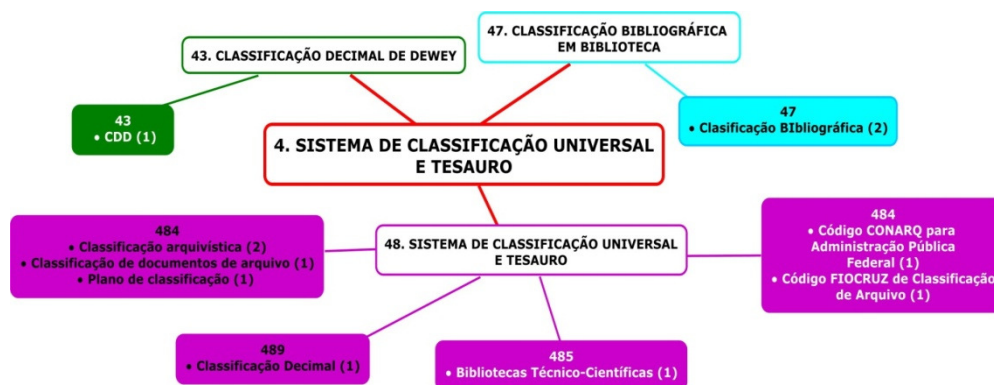
FIGURA 3: Classe 3 – Metodologia da Classificação e Indexação.



Fonte: Dados da pesquisa.

Na FIGURA 4, está representada a classe que enfoca as estruturas classificatórias universais e os tesouros, apresentando uma das menores ocorrências de termos: foram 9, no total, com 11 de frequência. Este dado demonstra que o campo tem pesquisado pouco a respeito de classificações universais. Tendo em vista a cobertura temporal da presente pesquisa (2003-2013), este pode ser um dado significativo, uma vez que tais estruturas continuam a ser amplamente utilizadas para a organização da informação em formato impresso. Gomes (2009) destaca a utilização destes instrumentos como apoio para a classificação automática e como fonte de vocabulário da linguagem natural para a indexação coordenada. Eles são igualmente interessantes para bases de dados gerais, servindo de apoio à organização da informação em sites.

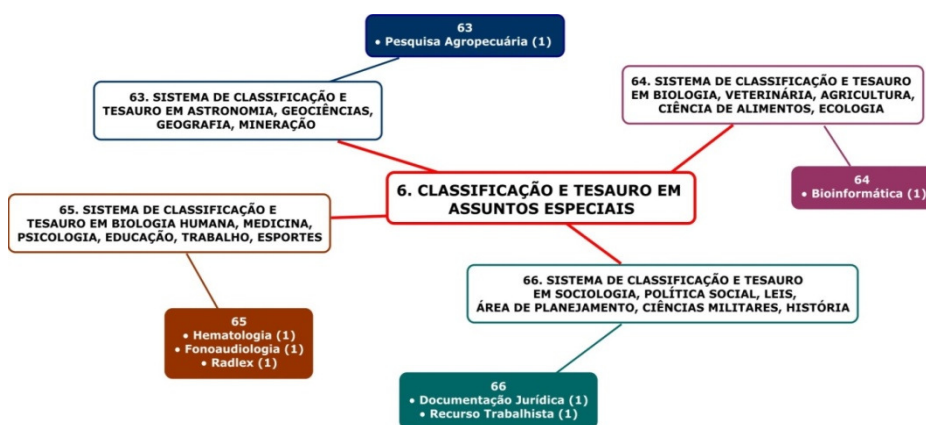
FIGURA 4: Classe 4 – Sistema de Classificação Universal e Tesouro



Fonte: Dados da pesquisa.

Na FIGURA 5, está representada a classe 6 com os termos que indicam a pesquisa aplicada em assuntos específicos, envolvendo uma abordagem voltada para os instrumentos classificatórios. As áreas com maior representação são aquelas relacionadas à Área Médica e Jurídica, nas quais foram classificados 7 termos com 7 frequências.

FIGURA 5: Classe 6 – Classificação e Tesouro em assuntos especiais.



Fonte: Dados da pesquisa.

A FIGURA 6 representa a classe 7, “Representação do Conhecimento por linguagem e Terminologia”, com 110 frequências em 47 termos. A subclasse 75 – “Tecnologias e Sistema de Recuperação On-Line” foi bem representada, com 40 de frequência em 18 termos. Este dado denota a preocupação da área com questões relativas à recuperação da informação digital. A subclasse 71 – “Problemas gerais da Linguagem Natural em relação à Organização do Conhecimento” – também apresentou diversidade de temas, com 20 de frequência em 8 termos. Com base nestes dados, entende-se que, os estudos do domínio estão priorizando a problemática dos aspectos linguísticos relacionados ao processamento da linguagem natural para representação e recuperação da informação.

FIGURA 6: Classe 7 – Representação do conhecimento por linguagem e Terminologia.



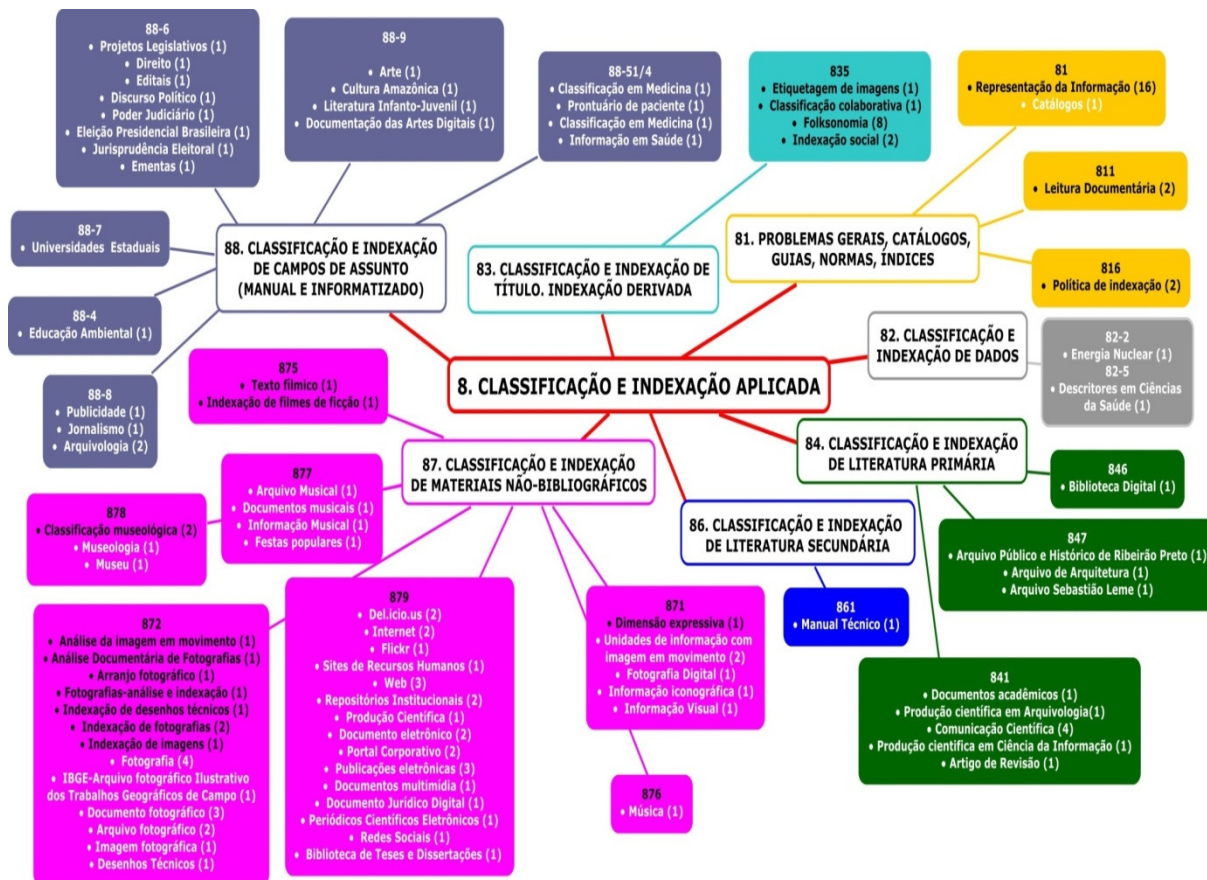
Fonte: Dados da pesquisa.

A FIGURA 7 representa a classe 8, destacando os métodos aplicados nas formas dos documentos e conteúdos de assunto no campo (DAHLBERG, 2006, p. 14). Esta classe teve 135 de frequência em 88 termos, sendo aquela com maior frequência e maior representatividade temática, o que denota a valorização da pesquisa aplicada na área. Nota-se ainda maior representatividade da classe 87, “Classificação e Indexação de materiais não bibliográficos”, com 43 termos e 62 frequências. Este dado deixa entrever a preferência em se fazer pesquisa aplicada, de forma mais concentrada, com objetos imagéticos e com aqueles relacionados ao contexto digital.

A classe 88, “Classificação e Indexação de áreas de assunto (manual e informatizada)” também foi bem representada, contando com 21 termos e 22 frequências. Nesta classe destacam-se os termos relacionados às Ciências Médicas e àqueles que tratam da área Jurídica.

Ao se observar o número de termos e o número de frequência da classe 8, percebe-se que, a maioria dos assuntos não chegou a ser abordada duas vezes, o que pode indicar falta de continuidade no estudo do objeto empírico - com a ressalva de que podem não ter sido representados por meio de palavras-chave ou serem relativos a pesquisas mais recentes.

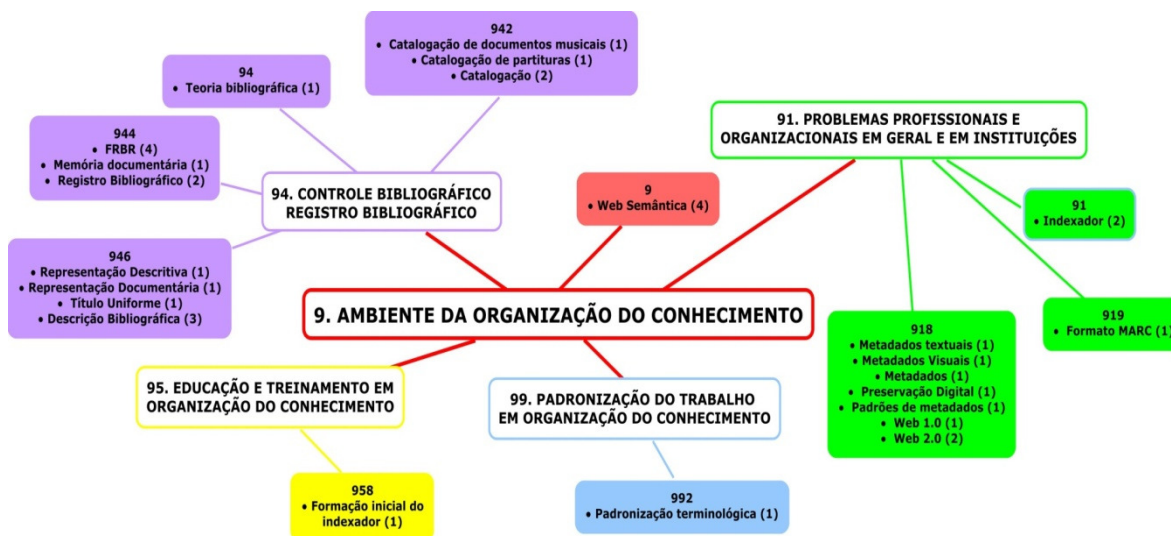
FIGURA 7: Classe 8 – Classificação e Indexação aplicada.



Fonte: Dados da pesquisa.

A FIGURA 8 nos mostra a classe 9, representando termos que compreendem os espaços influenciadores nas atividades da área, como questões de educação, leis, aspectos econômicos, estudos de usuários, padronização etc. (DAHLBERG, 2006, p. 14). Esta classe apresentou 23 termos com 35 frequências. A subclasse 94 teve maior variedade de temas: foram 11 termos com 18 de frequência. Os termos classificados nesta classe revelam as pesquisas na área de Representação Descritiva da Informação. Há de se destacar a pequena frequência de pesquisa nesta temática, tendo em vista a importância do tema para a área de Organização do Conhecimento.

FIGURA 8: Classe 9 – Ambiente da Organização do Conhecimento



Fonte: Dados da pesquisa.

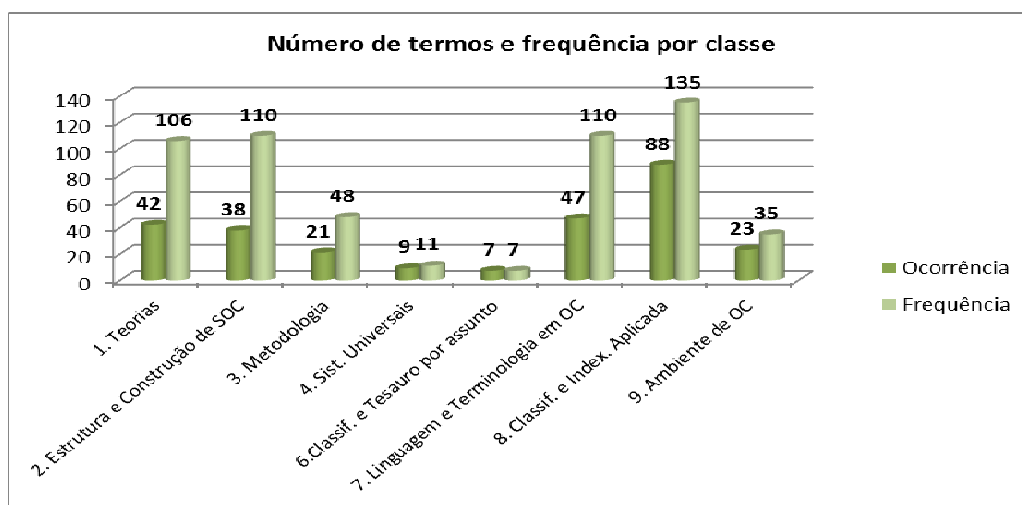
7 MAPEAMENTO TEMÁTICO: ANÁLISE E INTERPRETAÇÃO DOS DADOS

Os termos classificados e visualizados nos mapas demonstram as especificidades dos assuntos tratados em cada classe, e o modo como foram dispostos em cada subclasse. Buscou-se enquadrá-los do modo mais aproximado possível, pois, nem sempre a classe alcançou a especificidade do termo ou vice-versa. O principal objetivo da análise foi o de demonstrar o estado da arte, ou seja, de que modo os assuntos tratados no GT2 são representados na tabela, tendo em vista que nela são classificados diversos tipos de documentos da área de Organização do Conhecimento.

Foram classificados 275 termos, no total, não tendo havido recepção na tabela para os seguintes termos: Acumulação, Arquitetura da Informação, Bias, Gênese Documental, Identificação Arquivística, Leitura de Imagens, Memória Social, Memória, Respeito ao Fundo, Reuso de Vocabulários, Serviço Brasileiro de Respostas Técnicas, Tabela de Temporalidade e Tipologia Documental. Nota-se que vários destes termos referem-se à área da Arquivologia, o que demonstra não haver recepção na tabela para os termos desta área.

A observação do modo como as temáticas estão classificadas possibilitou a percepção de quais classes apresentam maior número de termos e maior frequência. Estes dados podem ser visualizados no GRÁFICO 1.

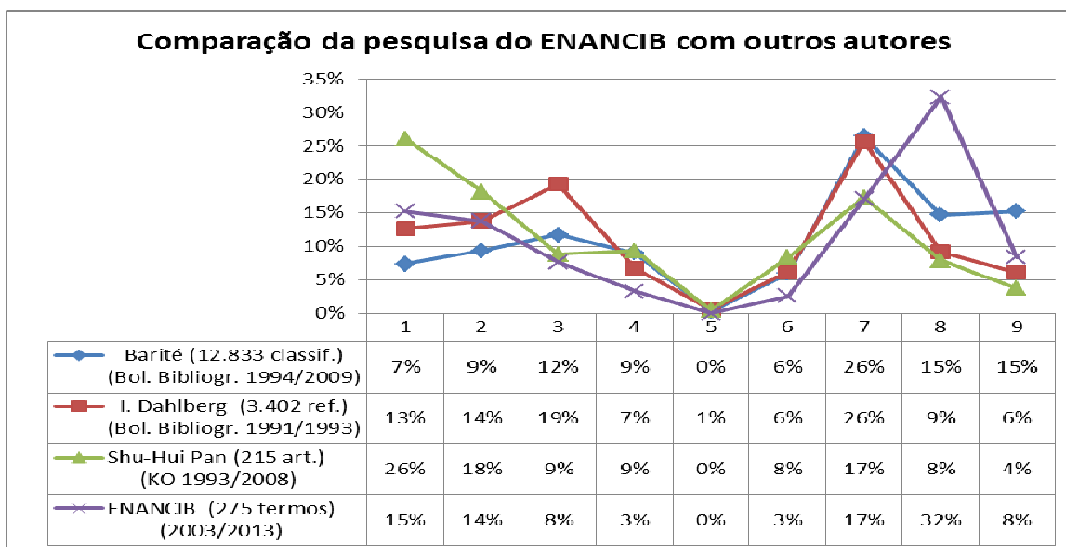
GRÁFICO 1: Número de termos e frequência por classe



Fonte: Dados da pesquisa.

Em uma análise da distribuição dos termos entre as 9 classes, percebeu-se que a interpretação seria mais rica se os dados fossem comparados com os de outras pesquisas, nas quais fosse possível identificar os aspectos distintivos do domínio. Sendo assim, buscaram-se pesquisas que tivessem classificado as temáticas da área com o mesmo modelo classificatório utilizado para os dados do ENANCIB. Foram identificadas três: Shu-Hui Pan (2010) publicou um *paper* sobre pesquisas na área de Organização do Conhecimento no qual analisou e classificou, pelo assunto principal, 215 artigos publicados no periódico “*Knowledge Organization*” no período de 1993 a 2008; Dahlberg (1995) utilizou esta tabela para sistematizar de que modo o domínio da Organização do Conhecimento vem trabalhando conceitualmente sua produção científica. A classificação dos trabalhos foi feita a partir das referências bibliográficas dos Boletins da “Literatura de Organização do Conhecimento”, editados pela ISKO no período de 1991 a 1993. Em sua tese de doutorado, Barité (2011) repertoriou as referências bibliográficas do mesmo boletim classificado por Dahlberg (1995), mas no período de 1994 a 2009. No gráfico 2, apresenta-se uma comparação entre as 4 pesquisas.

GRÁFICO 2: Comparação da pesquisa do ENANCIB com a de outros autores



Fonte: Dados da pesquisa.

Ao se analisar o gráfico 2, podem ser percebidas diferenças importantes entre as quatro pesquisas. Começando pela classe 1, “Fundamentos teóricos e problemas gerais”, nota-se que os aspectos teóricos são mais estudados nas pesquisas do GT2 do que naquelas representadas no Boletim Bibliográfico. No periódico *Knowledge Organization*, por sua vez, os estudos teóricos são priorizados e valorizados. Trata-se da classe com maior frequência.

Percebe-se certa regularidade nos percentuais da classe 4, “Sistema de Classificação Universal e Tesouro”, da classe 5, “Classificação de Objetos Especiais (*Special Objects Classifications - KOS*)”, da classe 6 “Classificação e Tesouro em Assuntos especiais” e da classe 9 “Ambiente de Organização do Conhecimento”, o que indica que, diante das demais classes, estas são as que apresentam menor representatividade nas quatro pesquisas. Este dado demonstra que as pesquisas do ENANCIB seguem uma tendência observada para os demais tipos documentais, o que nos leva a crer que se trate de uma tendência da área como um todo. Este dado carece de mais pesquisas, que deveriam analisar as causas para o desinteresse por estas temáticas e quais seriam as consequências para o domínio como um todo.

Nessa comparação, percebe-se que a classe 3 “Metodologia da Classificação e Indexação”, na pesquisa do ENANCIB, apresentou o menor percentual diante das demais pesquisas nesta mesma classe. Esta informação evidencia a necessidade de maior desenvolvimento das temáticas referentes aos métodos utilizados no domínio.

A classe 7, “Representação do Conhecimento por Linguagem e Terminologia”, é a mais pesquisada no Boletim Bibliográfico. Como este último reúne as referências de trabalhos sobre Organização do Conhecimento de forma mais abrangente, incluindo estudos de outros

domínios - além da Ciência da Informação - entende-se que tal temática desperta o interesse das demais áreas. Este dado apontaria para uma maior preocupação das demais áreas com questões referentes aos aspectos semânticos, ao processamento de linguagem automatizada, a tecnologias e sistemas de recuperação online e a problemas relacionados à Terminologia e à Representação do Conhecimento? Se a resposta é afirmativa, os dados reunidos no gráfico 10 comprovam que os aspectos teóricos - representados na classe 1 por esta mesma pesquisa - não estão sendo considerados, pois tiveram a menor frequência na comparação entre as quatro pesquisas analisadas. Se esta suposição for verdadeira, os dados comprovam uma alta incidência de estudos envolvendo as temáticas relacionadas à representação da informação por linguagem e terminologia com pequeno referencial teórico.

O aspecto mais característico das pesquisas do ENANCIB restringe-se à classe 8 “Classificação e Indexação Aplicada”, o quantitativo de pesquisa aplicada sendo muito superior às demais pesquisas, prova de que o domínio tem valorizado e acolhido esse tipo de pesquisa.

Tendo em vista a análise da distribuição das temáticas apresentadas no GT2 nas nove classes do modelo classificatório - e sendo a maioria destas comunicações fruto do trabalho dos Programas de Pós-Graduação do país - pode-se afirmar que os dados obtidos constituem um reflexo deste trabalho. Ainda assim, pesquisas mais detalhadas sobre os aspectos abordados se fazem necessárias, podendo trazer maior entendimento sobre os dados em questão.

8 CONSIDERAÇÕES FINAIS

O mapeamento temático de domínios do conhecimento evidencia os assuntos, os termos e os conceitos relacionados a um determinado saber, utilizando-se dos enunciados comunicativos das principais fontes de informação para mostrar os delineamentos que este último tem alcançado em sua área de atuação. Sendo a Ciência da Informação a área que estuda os processos informacionais que envolvem a representação e a organização da informação e do conhecimento visando uma recuperação precisa, os estudos atualizados de mapeamento temático revestem-se de vital importância para a área.

Esta pesquisa analisou as temáticas apresentadas no GT2, resultado das pesquisas desenvolvidas nos Programas de Pós-Graduação do país, visando entender de que modo o domínio tem se estruturado conceitualmente. Ainda que se reconheça a importância destes dados, entende-se que carecem de mais pesquisas de modo a ampliar as considerações aqui apresentadas. Sendo assim, fica a proposta para pesquisas futuras, que deveriam fornecer

respostas aos seguintes questionamentos: quanto de pesquisa com escopo em Organização do Conhecimento é apresentado fora do GT2, ou seja, em outros GTs? Quanto destas pesquisas é produzido fora do âmbito da Ciência da Informação? De que modo as pesquisas no domínio são representadas nas Teses e Dissertações e nos artigos de periódicos e outros eventos da Ciência da Informação? Por que o domínio menospreza a pesquisa referente à Representação Descritiva da Informação? Qual o motivo da pequena incidência de pesquisa nos aspectos metodológicos? A pesquisa que tem como fonte de dados o GT2 constitui uma pequena parcela de um vasto domínio.

REFERÊNCIAS

- BORBA, Diego dos Santos; VAN DER LAAN, Regina Helena; ROS CHINI, Bernadete. Palavras-chave: convergência e diferenciações entre a linguagem natural e a terminologia. **Perspectivas em Ciência da Informação**, v. 17, n. 2, abril/junho 2012.
- CHERNYI, A. I. On the problems of organization and representation of knowledge. **International Forum on Information and Documentation**, [S. l.], v. 22, n. 4, p. 3-10, 1997.
- DAHLBERG, I. Current trends in Knowledge Organization. In: GARCIA MARCO, F. J. Organización del conocimiento en sistemas de información y documentación 1. **Actas del I Encuentro de ISKO-España**, 1993, Madrid. Zaragoza: Librería General, 1995. p. 7-25.
- DAHLBERG, I. Fundamentos teórico-conceituais da classificação. **Revista de Biblioteconomia de Brasília**, Brasília, v.6, n.1, p.9-21, jan./jul. 1978.
- DAHLBERG, I. How to improve ISKO's standing: ten desiderata for knowledge organization, **Knowledge Organization**, Frankfurt, v. 38, n. 1, 2011.
- DAHLBERG, I. Knowledge organization: its scope and possibilities. **Knowledge Organization**, [S.l.], v.20, n.4, p. 211-222, 1993.
- FUJITA, M. S. L. Organização e Representação do Conhecimento no Brasil: análise de aspectos conceituais e da produção científica do ENANCIB no período de 2005 a 2007. **Tendências da pesquisa brasileira em Ciência da Informação**, v. 1, n. 1, 2008. Disponível em: <<http://inseer.ibict.br/ancib/index.php/tpbci/article/viewArticle/4>>. Acesso em: 15 dez. 2011.
- GOMES, Hagar Espanha. Tendências da pesquisa em Organização do Conhecimento. **Pesquisa brasileira em Ciência da Informação**, v. 2, n. 1, p. 60-88, jan./dez. 2009.
- GONZÁLEZ DE GÓMEZ, Maria Nérida. Para uma reflexão epistemológica acerca da Ciência da Informação. **Perspectiva em Ciência da Informação**, v. 6, n. 1, p. 5-18, jan./jun. 2001.
- GOPINATH, M. A.; DAS, P. Classification and representation of knowledge. **Documentation and Information Studies**, v. 34, n. 2, p. 85-90, 1997.
- GUIMARÃES, J. A. C. **Ensino de tratamento temático da informação nos cursos de biblioteconomia do Mercosul**: análise e perspectivas de um core curriculum à luz dos

avanços teóricos da área de Organização do Conhecimento. Marília: Faculdade de Filosofia e Ciências da UNESP, 2001.

HJORLAND, B. Domain analysis in information science. Eleven approaches - traditional as well as innovative. **Journal of the American Society for Documentation**, v. 58, n. 4, 2002. Disponível em:

<http://web.archive.org/web/20040721022850/http://www.db.dk/bh/publikationer/Filer/JDOC_2002_Eleven_approaches.pdf>. Acesso em: set. 2012.

HJORLAND, B. Domain analysis: a socio-cognitive orientation for Information Science research. **Bulletin of the American for Information Science and Technology**, v. 30, n. 3, p. 17-21, 2004.

HJORLAND, B.; ALBRECHTSEN, H. Toward a new horizon in information science: domain-analysis. *Journal of the American Society for Information Science*. v. 46, n. 2, p. 400-425, 1995. Disponível em:

<<http://comminfo.rutgers.edu/~kantor/601/Readings2004/Week3/r5.PDF>>. Acesso em: 5 maio 2012.

HJORLAND, Birger. Domain Analysis in Information Science. In: DRAKE, Miriam A. (ed.) **Encyclopedia of Library and Information Science**. 2. ed. Boca Raton: Taylor & Francis Group, 2005b. Disponível em: <<http://www.tandfonline.com/doi/book/10.1081/E-ELIS>>. Acesso em: 9 set. 2012.

HJORLAND, Birger; PEDERSEN, Karsten Nissen. A substantive theory of classification for information retrieval. **Journal of Documentation**, v. 61, n. 5, p. 582-597, 2005a.

MARTELETO, R. M. A pesquisa em Ciência da Informação no Brasil: marcos institucionais, cenários e perspectivas. **Perspectivas em Ciência da Informação**, v. 14, núm. esp., 2009.

SAYÃO, L. F. Modelos teóricos em Ciência da Informação – abstração e método científico, Brasília, **Ciência da Informação**, v. 30, n. 1, p. 82-91, jan./abr. 2001. Disponível em: <http://www.scielo.br/scielo.php?pid=S0100-19652001000100010&script=sci_arttext>. Acesso em: 5 dez. 2013.

SOUZA, R. F. de. Universo de ciência e tecnologia: organização e representação em classificações do conhecimento. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 5, n. 1, 2012. Disponível em:

<<http://inseer.ibict.br/ancib/index.php/tpbci/article/viewArticle/65>>. Acesso em: 16 jun. 2013.

**DA MÚSICA À ORGANIZAÇÃO DO ACERVO: UMA ANÁLISE DA
REPRESENTAÇÃO DA INFORMAÇÃO DA BIBLIOTECA JOSÉ ALBERTO
KAPLAN**

**FROM MUSIC TO THE COLLECTION'S ORGANIZATION AN ANALYSIS ON
INFORMATION'S REPRESENTATION OF THE JOSÉ ALBERTO KAPLAN'S
LIBRARY**

Ana Claudia Medeiros de Sousa
Maria Elizabeth Baltar Carneiro de Albuquerque

Resumo: A pesquisa objetivou analisar o tratamento, organização e recuperação dos itens documentais pertencentes ao acervo da Biblioteca José Alberto Kaplan, do Departamento de Música da Universidade Federal da Paraíba – Campus I, sob os parâmetros da representação da informação. Evidenciou-se que o tratamento e organização do acervo em questão são feitos de maneira rudimentar, uma vez que, constatou-se que apenas alguns livros receberam uma classificação a partir da Classificação Decimal Universal, já os demais itens documentais receberam um número de registro de acordo com que foi sendo adquirido. Concluiu-se que o nível de organização, armazenamento, representação dos itens documentais do acervo da Biblioteca Kaplan, não corresponde às práticas biblioteconômicas. Por fim, a pesquisa apresenta sugestões para efetiva organização, tratamento e representação da informação musical para referida Biblioteca.

Palavras-chave: Acervos musicais. Representação da Informação. Documentos musicais.

Abstract: The research aimed the treatment, organization and repair of items documentaries belonging to the collection of José Alberto Kaplan Library, of Department of Music of Universidade Federal da Paraíba – Campus I, underneath the parameters of representation of information. Was showed that the treatment and organization of this collection are made in a primitive way, once was shoed that only some books received a rating from Universal Decimal Classification, already the others documentaries items receveid just one number of register according with was being acquired. Was understood that the standard of organization, storage and representantion of documentaries items of collection from the Kaplan Library does not corresponds to the exercise of the librarianship. Lastly, the research presents suggestions to the effective organization, treatment and representation of the musical information of the cited Library.

Keywords: Musical Collection. Representation of Information. Musical Documentaries.

1 INTRODUÇÃO

A informação faz parte do cotidiano de qualquer pessoa e a sua importância vai depender da necessidade de cada um. Assim, a informação se torna indispensável para o indivíduo, mas para obter um bom resultado da sua utilização é imprescindível que haja um tratamento da mesma. Um dos desafios e objetivos da Ciência da Informação é investigar os processos relacionados ao tratamento, organização, representação, recuperação, acesso e uso da informação.

Com isso, a representação da informação é uma das bases da Ciência da Informação, Biblioteconomia e Arquivologia, pois é a partir dela que os pesquisadores e profissionais das

áreas citadas representam o conteúdo de documentos, proporcionando aos usuários uma prévia análise do teor do documento, facilitando o processo de recuperação da informação. Sendo assim, a representação é considerada uma atividade determinante para a recuperação de informações contida nos mais variados documentos. A representação da informação em sua essência apresenta previamente aos usuários, as temáticas, os conteúdos abordados nos documentos.

Assim, os mais diversos tipos de acervos documentais fazem uso da representação da informação, para proporcionar a organização, armazenamento, recuperação e uso da informação. Com os acervos de documentos musicais não é diferente, daí a importância dos Musicólogos, Cientistas da Informação, Bibliotecários e Arquivistas trabalharem em conjunto para proporcionar um tratamento adequado aos acervos musicais, até mesmo porque como cita McGarry (1999, p. 11) “a informação deve ser ordenada, estruturada ou contida de alguma forma, senão permanecerá amorfa e inutilizável”. Com isso, a informação necessita receber um tratamento para que possa ser recuperada.

A temática ‘Acervos Musicais’, tem sido de interesse de alguns estudiosos, como Cotta (2006), Castagna (1997), Freire (2002), cujas pesquisas trilham na perspectiva da memória da música brasileira e patrimônio cultural.

Já para os estudos relacionados à Representação da Informação de documentos musicais, pode-se citar o trabalho de Souza (2008), em que o mesmo apresentou uma sugestão de um Microtesauro em Música, intitulado ‘MiMu’. O referido tesouro utiliza a seguinte legenda:

CAT: Categoria (Autoridades, Equipamentos, Gêneros e estilos, Saúde, Síntese sonora, Sociedade e Teoria musical);
NE: Nota explicativa. Inclui a definição do termo ou dica de uso;
USE: Remissiva Ver, utilizada para indicar o descritor autorizado;
UP: Remissiva Usado por, utilizada para indicar os não-descritores;
TG: Termo geral, ou genérico, numa relação hierárquica;
TE: Termo específico, numa relação hierárquica, subordinado ao termo geral;
TR: Termo relacionado, para uma relação associativa (Souza, 2008, p. 111) (grifo nosso).

O MiMu foi estruturado com termos baseados em artigos da revista Teclado & Áudio e do Dicionário Grove de Música, devido a especialização dos assuntos tratados na área da música. Com isso, o tesouro segue a seguinte estrutura:

Acústica
CAT: SÍNTESE SONORA

NE: "A ciência do som e da audição. Trata das qualidades sônicas de recintos e de edificações, e da transmissão do som pela voz, por instrumentos musicais ou por meios eletrônicos." (DICIONÁRIO GROVE, 1994, p. 6)

TG: Som

TE: Tratamento acústico

TR: Amplificação
Sonorização

Adagio

CAT: TEORIA MUSICAL

NE: "(do italiano adagio, 'à vontade', 'calmamente') Um movimento em andamento lento." (DICIONÁRIO GROVE, 1994, p. 6)

TG: Andamento

TR: Andante

Grave

Larghetto

Largo

Lento (Souza, 2008, p. 117) (grifo nosso).

A justificativa de Souza (2008) para desenvolver o microtesauro foi à pretensão de melhorar os processos de indexação e recuperação da informação musical, percebe-se pela maneira em que foram elaborados os termos, que o autor atingiu o objetivo proposto no trabalho, com a elaboração de um tesauro especializado para acervos musicais.

Tanto os acervos musicais, como arquivos de orquestras enfrentam problemas relacionados à organização, tratamento e representação de documentos musicais e partituras. Para tanto, este estudo busca evidenciar a importância da representação da informação para organização de acervos musicais, a partir da análise da representação.

Destarte, esta pesquisa teve o objetivo de diagnosticar o tratamento, organização e recuperação de itens documentais pertencentes ao acervo da Biblioteca José Alberto Kaplan, do Departamento de Música da Universidade Federal da Paraíba – Campus I, sob a ótica da representação da informação. A pesquisa propõe ainda, sugestões para a criação de sistemas de representação da informação para propiciar uma efetiva recuperação e uso da informação.

Para buscar os dados e alcançar resultados da pesquisa, fez-se necessária a adoção de métodos e técnicas que direcionaram o processo de investigação. Nesta perspectiva, trata-se de uma pesquisa exploratória, uma vez que, este tipo de pesquisa “se caracteriza pelo desenvolvimento e esclarecimento de ideias, com o objetivo de oferecer uma visão panorâmica, uma primeira aproximação a um determinado fenômeno que é pouco explorado” (GONSALVES, 2007, p. 67). Inicialmente houve uma aproximação do acervo musical em questão, com o intuito de analisar a representação dos itens do acervo e em seguida apresenta

sugestões para representação da informação dos materiais pertencentes à Biblioteca de Música José Kaplan.

2 REPRESENTAÇÃO DA INFORMAÇÃO

A Ciência da Informação investiga os processos relacionados à informação, desde a sua produção até a disponibilização, atuando efetivamente no fluxo informacional, com intuito de propiciar os meios para otimizar o acesso e uso, para atender as necessidades informacionais dos usuários. A produção desenfreada de informação científica despertou uma inquietação com o tratamento, representação e recuperação destas produções.

Assim, a análise de princípios, meios e técnicas que norteiam a representação da informação tem despertado cada vez mais interesse por parte dos Cientistas da Informação, Bibliotecários e Arquivistas, uma vez que, a evolução dos suportes informacionais tem tornado-se comum, proporcionando novos desafios nas formas de tratamento da informação. Com isso, a representação da informação configura-se como essencial para organização e recuperação da informação. Alvarenga (2003, p. 4) cita que o processo de representação

[...] possui as etapas de percepção, identificação, interpretação, reflexão e codificação, etapas que são envolvidas no ato de se conhecer um novo ser ou uma coisa, ou aprofundar-se no conhecimento de um ser ou uma coisa já conhecida, utilizando-se dos sentidos, da emoção, da razão e da linguagem.

Para McGarry (1999), o processo de representação da informação, deve adotar alguma forma de veículo, quer dizer, deve possuir atributos essenciais como, sinais, signos e símbolos, os quais contribuem e atuam no processo de representação e compreensão de profissionais da informação e principalmente dos usuários. Furgeri (2006, p. 27) descreve que

Os **sinais** estabelecem relações com as ações a serem desenvolvidas pelo receptor. [...] Os **signos** indicam a presença física de algo ou algum evento relacionado a eles. [...] Os **símbolos** tendem a possuir significados mais duradouros e constituem-se em representações culturalmente construídas e reconhecidas por uma comunidade específica. (grifo nosso)

Além dos três veículos apresentados anteriormente, McGarry (1999) cita ainda a linguagem, considerando-a como fundamental para comunicação humana. Para Novellino (1996) as ações de comunicação humana, as quais têm a linguagem como instrumento, podem se realizar mediante falante e ouvinte, ou entre a imagem e aquele que assiste, ou ainda entre texto e leitor. Esta ação comunicativa pode ser abordada de duas maneiras, sob o ponto de vista da recuperação da informação ou da representação da informação. A autora frisa também que a principal característica do processo de representação da informação é a

[...] substituição de uma entidade lingüística longa e complexa - o texto do documento - por sua descrição abreviada. O uso de tal sumarização não é apenas uma consequência de restrições práticas quanto ao volume de

material a ser armazenado e recuperado. Essa sumarização é desejável pois sua função é demonstrar a essência do documento (NOVELLINO, 1996, p. 38).

Assim, a representação da informação atua como um artifício para destacar o que é fundamental no documento, a partir do uso de símbolos, signos e sinais, os quais possibilitam a recuperação do conteúdo abordado nos documentos. De acordo com Novellino (1996, p. 38), o processo de Representação da Informação envolve dois itens a ser destacados:

1) análise de assunto de um documento e a colocação do resultado desta análise numa expressão lingüística.

2) atribuição de conceitos ao documento analisado.

Na atribuição de conceitos faz-se necessária a adoção de instrumentos de padronização, “a qual visa garantir que indexadores de um mesmo sistema ou sistemas afins usem os mesmos conceitos para representar documentos semelhantes”. (NOVELLINO, 1996, p. 38)

A representação da informação requer uma importante prática que é a análise da informação, a qual se preocupa com o conteúdo do documento. A análise da informação viabiliza - tanto por modelos conceituais ou técnicas e ferramentas da documentação, a efetiva representação da informação.

Para tanto, os profissionais da informação utilizam instrumentos que auxiliam no processo de representação, tais como tabelas de classificação, cabeçalhos de assuntos, tesouros, entre outros, na tentativa de representar o conteúdo informacional de maneira eficiente. Tratam-se de linguagens documentárias, que são ferramentas que possuem conjuntos de termos, símbolos e signos controlados, os quais são utilizados para representar informações contidas nos documentos. O objetivo das linguagens documentárias é padronizar o vocabulário utilizado para representar informações.

As linguagens documentárias foram inicialmente utilizadas com o intuito de padronizar entradas de assunto de catálogos, as listas de cabeçalho de assuntos, a partir da utilização das Linguagens documentárias, percebeu-se que estas seriam ferramentas de indexação, representação e recuperação da informação.

Campos (1995, p. 1) cita que, os “instrumentos como a tabela de classificação, o tesouro e as terminologias, sistematizam os conceitos de uma área de conhecimento na perspectiva de representar e possibilitar a recuperação das informações”.

Portanto, a representação é uma atividade indispensável para o tratamento da informação, uma vez que, o processo de representação torna-se determinante para organização, recuperação, acesso e uso do suporte informacional.

3 REPRESENTAÇÃO DE ACERVOS DE DOCUMENTOS MUSICAIS

Os registros e linguagem da música foram desenvolvidos ao longo do tempo. Os primeiros registros musicais foram através de sinais manuais e do primeiro sistema alfabético grego, datados do século 500 a.C., já pelos chineses os primeiros registros musicais são do século 3 a.C. e os textos hebraicos surgiram no século VI. Souza (2008, p. 18) cita quanto aos registros musicais que

[...] na cultura ocidental reinicia no IX e dois séculos depois nas igrejas orientais; por fim no século XI Guido d'Arezzo é apontado como um dos responsáveis pela idealização da pauta moderna, notação musical que sofreu importantes modificações até o decorrer do século XX.

Quanto à guarda e preservação dos registros musicais, bibliotecários, arquivistas e documentalistas aparecem como responsáveis pela conservação destes documentos criados ao longo da história, o que possibilitou a concentração de coleções de obras artísticas, como as partituras, as quais atravessaram o tempo, tornando-se algumas de domínio público, outras em obras raras (SOUZA, 2008).

Atualmente, os acervos musicais têm aliado conceitos e técnicas da biblioteconomia e arquivística integrada às necessidades específicas para o tratamento técnico de acervos ligados à música. Este fato tem despertado interesse por parte de pesquisadores da área de representação, descrição e recuperação da informação, como afirmam Sena e Alves (2012, p. 2)

A questão da representação e recuperação da música aparece como objeto de estudo da Ciência da Informação pela primeira vez em 1996, é um tema de pesquisa que cresceu recentemente com a explosão do interesse de coleções em rede; porém, pesquisas nessa área, no Brasil, são escassas na literatura da Ciência da Informação.

A Sociedade Internacional de Musicologia em parceria com a Associação Internacional de Bibliotecas, Arquivos e Centros de Documentação de Música, foram os pioneiros em desenvolver normas de catalogação de documentos musicais. Tal iniciativa ocorreu na década de 1950 (SENA; ALVES, 2012)

Apesar de atualmente existir algumas ferramentas que auxiliam no processo de descrição e representação de documentos musicais, ainda há muitas dificuldades por parte dos profissionais da informação, em tratar tais documentos, por não compreender determinadas peculiaridades da música.

De acordo com o Dicionário Grove (1994, p. 656) a notação musical é “um equivalente visual do som musical, que se pretende um registro do som ouvido ou imaginado, ou um conjunto de instruções visuais para interpretes”. Portanto, o registro informacional

musical, representa através de linguagens: os símbolos e signos, as notas, pausas, dinâmicas, descrições etc., que devem ser executadas pelo intérprete/músico.

Daí a importância do profissional da informação compreender a linguagem musical, entender os conceitos como, por exemplo, da pauta musical, clave, movimentos, tonalidades, estilos musicais, séries etc. Sena e Alves (2012, p. 3) conceituam pauta musical como

[...] a estrutura usada para a notação musical, formada por um conjunto de cinco linhas paralelas e equidistantes formando entre si quatro espaços. As sete notas musicais são escritas tanto sobre as linhas, como nos espaços entre elas e se organizam em ordem gradual de altura. Para convencionar o posicionamento das notas na pauta é usado um sinal chamado “Clave” que se coloca no princípio da pauta. A clave fixa a altura de uma das cinco linhas da pauta, dando a orientação para o reconhecimento das outras linhas e espaços. (SENA; ALVES, 2012, p. 3)

Percebe-se assim, que o tratamento e organização da informação musical requerem alguns conhecimentos particulares da área da música, pois tanto os profissionais da informação encontram dificuldades em tratar a informação musical, como os musicólogos, que apesar de conhecer a linguagem musical, desconhecem das técnicas e ferramentas de descrição, representação e organização da informação. Como cita Faria (2009, p. 2)

[...] os bibliotecários e arquivistas geralmente não têm conhecimento musical suficiente para atender às necessidades informacionais dos músicos e regentes e estes, de modo geral, desconhecem técnicas e padrões biblioteconômicos ou arquivísticos estabelecidos de tratamento documental.

Para tanto, cabe aos profissionais da informação, juntamente com os da música buscar compreender a complexidade e particularidade dos acervos musicais e a partir de um trabalho conjunto, proporcionar o tratamento do acervo musical. Ou então o Bibliotecário/Arquivista terá que ter um mínimo de conhecimento da linguagem musical. Para Assunção (2005, p. 3)

[...] la documentación musical continua a ser mirada por los bibliotecarios y archiveros como un dominio hermético de músicos y musicólogos y las técnicas documentales siguen a ser miradas por los musicólogos como complicaciones inútiles de tecnócratas. Unos y otros están equivocados.

Assim, o tratamento e organização da informação musical tem se tornado de interesse de pesquisadores, o que tem resultado em algumas experiências e soluções para descrição e representação da informação musical, como por exemplo, o estudo de McLane (1996) que apresenta três visões:

Visão Subjetiva: o uso do esquema de notação para representar a obra musical pode ser considerado a visão subjetiva da obra. Subjetiva porque a escolha de elementos de notação normalmente representa uma obra em “contexto-dependente” no sentido de que a decisão da notação pode incluir ou excluir aspectos particulares da obra. A informação bibliográfica descritiva como parte da visão subjetiva também se inclui aqui.

Visão Objetiva: um som gravado pode ser identificado como uma visão objetiva da obra musical. O som musical é objetivo porque uma vez gravado,

a representação da música através da gravação é fixada e não mais sujeita as variações editoriais e de performance. Esta visão pode ser considerada a mais completa representação da música.

Visão Interpretativa: a representação através da análise de alguns aspectos da obra. Classificações e esquemas analíticos que ilustram características que não são óbvias de uma obra musical, como o gênero musical, ou de um conjunto de obras, entram nesta categoria.

Desta maneira, as visões de representação apresentadas por McLane são consideradas tanto complementares, como interdependentes para a recuperação da informação da música. De acordo com Sena e Alves (2012, p. 2) “a recuperação da informação da música depende tanto da complexidade quanto da forma como a informação é representada, e de um conhecimento prévio do usuário para encontrar a informação desejada”.

Pelo o fato de o usuário ter um conhecimento prévio da informação desejada, o profissional da informação também precisa ter algum conhecimento a respeito do documento musical. Desta maneira, para o tratamento e representação de tais documentos, além de considerar as práticas biblioteconômicas, faz-se necessário o profissional compreender as necessidades dos usuários e as maneiras que eles buscam a informação.

Independente do suporte informacional, a representação da informação propicia a recuperação do documento, sejam livros, partituras, CD's etc. Com isso, a representação da informação musical, tem se otimizado a partir da utilização de tesouros e sistemas de classificação, os quais têm contribuído no tratamento, organização, descrição e representação da informação.

4 A REPRESENTAÇÃO DA INFORMAÇÃO MUSICAL NA BIBLIOTECA JOSÉ ALBERTO KAPLAN

A Biblioteca José Alberto Kaplan foi criada no ano de 1978, período em que surgiu o curso de bacharelado em Música na Universidade Federal da Paraíba. Sabe-se que Bibliotecas que atendem as demandas de Instituições de Ensino Superior, devem estar inseridas no contexto do processo evolutivo das áreas do conhecimento as quais pertencem, com o intuito de contribuir, acima de tudo, para o desenvolvimento científico, tecnológico e social, a partir da disseminação da informação. Com essa proposta, as Bibliotecas tornam instrumentos de multiplicação de difusão cultural, promovendo a integração da comunidade acadêmica da Instituição.

Certamente este foi o intuito da criação da Biblioteca José Alberto Kaplan, atender especificamente ao curso de Bacharelado em Música da UFPB, como também aos profissionais da música da comunidade geral. Seu acervo é composto de aproximadamente 3.000 (três mil) títulos, formado de livros de teoria musical, história da música, partituras,

discos de vinil, CDs, entre outros. Tais itens foram adquiridos tanto por compra como também através de doações feitas pelos músicos, professores, compositores, alunos e ex-alunos do Departamento de Música da UFPB.

Os usuários da referida Biblioteca são discentes e docentes dos cursos de Bacharelado em Música, Licenciatura em Música, Sequencial em Música como também os da Pós-Graduação em Música da UFPB.

No que se refere ao tratamento e organização do acervo em questão, apenas alguns livros receberam uma classificação a partir da CDU, já os demais itens documentais receberam um número de registro de acordo com que foi sendo adquirido. Assim, aproximadamente 20% do acervo passou pelo tratamento de catalogação, descrição e classificação, os demais possuem um número de registro. Este fato é preocupante, pois se sabe da importância da padronização da descrição e representação da informação.

Já as partituras estão organizadas em ordem alfabética, por autor e por instrumento musical. O que é insuficiente para propiciar a efetiva representação e recuperação da informação musical.

Vale salientar que os itens que receberam tratamento, como catalogação, classificação e número de registro, foi em um período que o Departamento de Música da UFPB buscou parceria com a Biblioteca Central da referida universidade, na tentativa de organizar o acervo da Biblioteca Kaplan, a luz das práticas biblioteconômicas. Assim, foi uma Bibliotecária encarregada de ensinar aos demais funcionários, maneiras de tratamento, organização e representação da informação. Apesar de tal iniciativa, foi insuficiente, pois a maior parte do acervo continuou sem receber o devido tratamento.

Evidencia-se assim, que o nível de organização, armazenamento, representação dos itens documentais do acervo da Biblioteca Kaplan, não corresponde às práticas biblioteconômicas. Percebe-se com isso, a importância da atuação dos bibliotecários e musicólogos, no que se refere à organização de acervos musicais. Tais profissionais devem compreender alguns elementos, que segundo Sena e Alves (2012, p. 7) são de extrema importância para os usuários de acervos musicais. Tais elementos são determinantes para recuperação da informação musical. São eles:

Dados sobre Compositor:

Nome: indicação do nome completo do compositor dando destaque ao nome usual, bem como a nomes alternativos, pseudônimos.

Nascimento: data e local de nascimento.

Residência: muitos são os compositores que deixam sua terra natal e migram para outros países, por terem desenvolvido atividades significativas em outro país, é importante ter a recuperação desse local.

Morte: data e local da morte.

Contexto no qual a obra foi composta:

Data da Composição: indicação da data em que a partitura foi concluída, ou do período em que foi composta.

Local de Composição: local onde a música foi concebida e a partitura escrita.

Data da Estréia: data exata ou aproximada na qual a obra foi executada publicamente pela primeira vez.

Local da Estréia: localidade onde ocorreu o evento da primeira execução. Evento da Estréia. Interpretes da Estréia.

Encomenda: indicação de quem fez a encomenda da obra.

Informações Técnicas: duração da obra, descrição física dos documentos, descrições de conteúdo.

Informações sobre época, estilo, estrutura e técnica de escritura: no que se refere a tendências, estilo e técnicas composicionais.

Elementos musicais explorados na obra: informações básicas sobre os aspectos composicionais da obra servem como orientação didático-musical aos professores de música, educadores musicais e regentes de corais, na escolha das obras para suas atividades pedagógicas.

Grau de dificuldade: é definido por uma escala simples com quatro níveis de dificuldade: Elementar, Médio, Avançado e Virtuoso.

Torna-se relevante ressaltar a riqueza de detalhes com que as autoras Sena e Alves (2012) descrevem as características e informações fundamentais para o tratamento da informação musical.

Além da ausência na padronização da representação e organização dos itens do acervo da Biblioteca Kaplan, a mesma ainda não é automatizada, com isso as técnicas de tratamento de informação utilizadas são manuais, inclusive o empréstimo.

A Biblioteca em questão não faz parte do Sistema de Bibliotecas da UFPB. Este sistema é formado por quinze bibliotecas, sendo uma central e quatorze setoriais, integradas sob os aspectos funcional e operacional. Assim, percebe-se que apesar da Biblioteca Kaplan pertencer a UFPB, a mesma não está inserida no sistema, não é automatizada e principalmente não possui um profissional bibliotecário. Este fato compromete seu acervo, uma vez que, sabe-se da importância da atuação do profissional bibliotecário em unidades de informação, pois

[...] o bibliotecário assume papel de co-educador, criando um diferencial perante usuários que necessitam e utilizam a informação. Este profissional tornará possível a recuperação e disseminação da informação armazenada nas unidades de informação, independentemente de sua localização física,

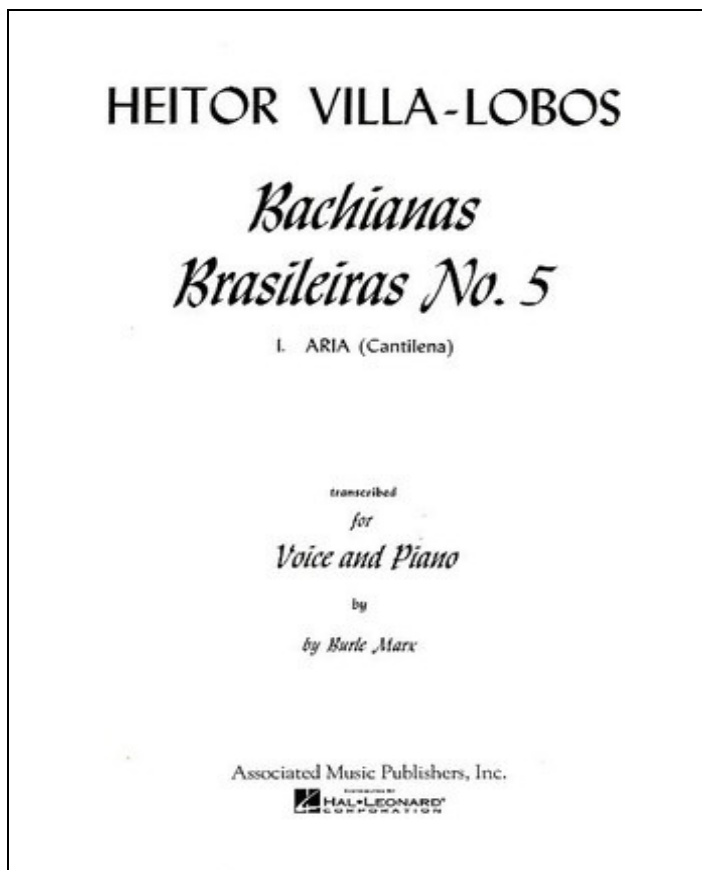
possibilitando assim a plena utilização das informações de forma interativa e dinâmica proporcionando o aprimoramento do saber e a otimização das necessidades informacionais dos usuários (SANTOS-ROCHA, s. d., p. 10)

Os funcionários que trabalham atualmente na Biblioteca Kaplan, além de não possuírem formação na área de Biblioteconomia, não possuem também na área de Música, talvez essa seja uma das justificativas para a falta de padronização na organização de seu acervo. Como o tratamento da informação é feito de maneira rudimentar, não é adotado sistemas de classificação, tesouros, listas de cabeçalhos de assuntos.

5 TRATAMENTO DA INFORMAÇÃO MUSICAL

Para o tratamento de catalogação, classificação e descrição de itens de acervos musicais, requer a utilização de instrumentos que viabilizem tal processo. Para tanto, será exemplificado a seguir, o tratamento de uma partitura, com o auxílio do Manual de Catalogação de Partituras da Biblioteca da ECA/USP, do Sistema de Classificação Universal (CDU), da Tabela de *Cutter*, e ainda o Microtesauro em Música (MiMu):

FIGURA 1: Partitura Bachianas Brasileiras No. 5 de Villa-Lobos



Fonte: <http://www.di-arezzo.es/partitura/partitura+c1%C3%A1sica/partitura-para-voz/SCHIR01603.html>

FIGURA 2: Grade da Partitura Bachianas Brasileiras No. 5 de Villa-Lobos

Aria (Cantilena)
(from *Bachianas Brasileiras No. 5*)
Composed by Heitor Villa-Lobos

©Copyright 1947 (Renewed) by Associated Music Publishers Inc., USA.
The arrangement ©Copyright 1993 by Associated Music Publishers Inc., USA.
All Rights Reserved. International Copyright Secured.

Preview at www.musicaneo.com

Fonte: <http://www.di-arezzo.es/partitura/partitura-cl%C3%A1sica/partitura-para-voz/SCHIR01603.html>

A partir da utilização do Manual de Catalogação de Partituras da Biblioteca da ECA/USP, que segue a categoria de informação indicada:

LOCALIZAÇÃO:
 AUTOR: VILLA-LOBOS, Heitor (1887-1959)
 TÍTULO: Bachianas No. 5: Ária Cantinela
 MEIO DE EXPRESSÃO: Voz e piano
 NOTAS: In: Homenagem a Sebastian Bach
 LOCAL: s. l.
 EDITORA: G. Schirmer Inc.
 DESCRIÇÃO FÍSICA: 12 p.
 PAÍS DO AUTOR: Brasil

Fonte: Manual de Catalogação de partituras da Biblioteca da ECA.

Com o auxílio da categoria de informação, dos descritores retirados do Microtesouro em Música (MiMu), *Cutter* e CDU, resultou na seguinte representação descritiva:

78.087.4 V712b [s. 1] : Ed. Erudita.	Villa-Lobos, Heitor. Bachianas brasileiras n. 5 : Ária Cantinela / Heitor Villa-Lobos; [Copista] Burle Marx. — G. Schirmer, 1938. 12 p. 1. Voz. 2. Piano. 3. Partitura. 4. Música I. Título. II. Burle Marx.
---	--

6 CONSIDERAÇÕES

A representação da informação influencia nos processos de tratamento, descrição e organização da informação. Assim, ela torna-se determinante para recuperação, atendendo as demandas informacionais de usuários.

O estudo buscou evidenciar a organização e representação da informação utilizada na Biblioteca José Alberto Kaplan, e ficou evidente que a referida biblioteca não possui padronização na representação e organização da informação.

As Bibliotecas devem oferecer uma infra-estrutura bibliográfica compatível a suas atividades, e está voltada a satisfação dos usuários. Dispor ainda de ambiente acolhedor, computadores que permitam acesso imediato às informações desejadas através de catálogos on-line de autor, título e assunto, entre outros.

Com isso, o estudo buscou levantar opções de sugestões para contribuir no processo de representação e organização da informação da Biblioteca Kaplan. Tais sugestões são:

- Providenciar a contratação de um profissional bibliotecário. Sabe-se que é fundamental a gerencia deste profissional em uma Biblioteca com acervo com mais de 200 itens documentais;
- A Biblioteca Kaplan passar a integrar o Sistema de Bibliotecas da UFPB, para se agregar ao planejamento, distribuição de recursos, padronização dos serviços etc. do Sistema citado;
- O acervo ser automatizado a partir do sistema de informação adotado na UFPB, o Sistema Integrado de Gerenciamento de Atividades Acadêmicas - SIGAA;
- Adotar as normas de catalogação de documentos musicais propostas pelo Manual de catalogação de partituras da Biblioteca da ECA/USP. A catalogação registra informações de um item documental, com a adoção de uma padronização para documentos musicais, haverá um tratamento e recuperação mais eficaz;
- Adotar ISBD (PM) – Norma Internacional de Descrição Bibliográfica para Música Impressa. Agregar esta norma com a citada anteriormente, para adequar as necessidades provenientes do tratamento de documentos musicais;
- Adotar o Tesouro de Música ‘MiMu’, para representação das temáticas abordadas nos itens do acervo com maior precisão;

- Receber alunos do curso de Biblioteconomia para as atividades de Práticas de Laboratório exigidas na Graduação. Oportunidade para os alunos praticarem atividades em acervos musicais.

Independente do suporte informacional, espécie documental, biblioteca especializada, biblioteca especial, enfim, seja qual for o tipo de informação a ser tratada, cabe ao profissional da informação, buscar técnicas e métodos da biblioteconomia e documentação para tratá-los de maneira que venha favorecer sua organização e principalmente a recuperação no momento oportuno. Garantir assim, o acesso à informação.

REFERÊNCIAS

- ALVARENGA, L. Representação do conhecimento em tempo e espaço digitais. *Encontros Bibli*, Florianópolis, n. 15, 2003.
- CAMPOS, M. L. A. Perspectivas para o estudo da área de Representação da Informação. *Ciência da Informação*. v. 25, n. 2, p. 1-7, 1995.
- COTTA, A. G. Fundamentos para uma arquivologia musical. In: COTTA, A. G.; BLANCO, P. S. *Arquivologia e patrimônio musical*. Salvador: Edufba, 2006.
- DICIONÁRIO *Grove de Música*. Rio de Janeiro: Jorge Zahar, 1994.
- Disponível em: <<http://pt.scribd.com/doc/109279048/Musica-como-informacao-criterios-para-catalogacao-de-partituras>>. Acesso em: 05 jul. 2013.
- FARIA, M. M. de. O tratamento documental dos arquivos musicais e a busca de práticas comuns no tratamento da música brasileira para orquestra. *Opus: Revista Eletrônica da ANPPOM*. Goiânia, v. 15, n. 1, p. 85-90, jun. 2009.
- FURGERI, S. *Representação de informação e conhecimento: estudo das diferentes abordagens entre a Ciência da Informação e a da Ciência da Computação*. 2006. Dissertação (Pós-Graduação em Ciência da Informação). Campinas, 2006.
- GONSALVES, E. P. *Conversas sobre iniciação à pesquisa científica*. Campinas: Alínea, 2007.
- McGARRY, K. *O contexto dinâmico da informação*. Brasília: Briquet Lemos, 1999.
- NOVELLINO, M. S. F. Instrumentos e metodologias de representação da informação. *Inf. Inf.*, Londrina, v.1, n.2, p.37-45, jul./dez. 1996.
- RECINE, A. S. V.; MACAMBYRA, M. *Manual de Catalogação de partituras da Biblioteca da ECA*. São Paulo: Serviço de Biblioteca e Documentação/ECA/USP, 2010.
- SANTOS-ROCHA, E. S. *O papel do bibliotecário como mediador no desenvolvimento da competência em informação na universidade*. Disponível em: <<http://pt.slideshare.net/marcrisfer/papel-do-bibliotecario>>. Acesso em 13 jul. 2013.
- SENA, A. C. M. de; ALVES, W. C. S. *Música como Informação: critérios para catalogação de partituras*. I ENACAT. 2012.

SOUZA, R. L. L. de. *Microtesouro em Música: teoria e prática*. 2008. 564f. Monografia (Departamento de Ciência da Informação da Universidade de Brasília), Brasília, 2008.

ORGANIZAÇÃO DOS ESTUDOS INTERDISCIPLINAR DE SUBSTÂNCIAS PSICOATIVAS

ORGANIZATION OF INTERDISCIPLINARY STUDIES OF SUBSTANCES PSYCHOACTIVE

Miriam Gontijo Moraes

Resumo: O objetivo desse artigo é apresentar a estratégia para a identificação do sistema nocional dos estudos de substâncias psicoativas visando à elaboração de uma proposta de Estrutura Conceitual para um domínio interdisciplinar. A estratégia adotada refere-se aos instrumentos e metodologias de construção de sistemas de organização do conhecimento e modelagem conceitual, em especial, os relacionados à elaboração de tesouros. Além da Análise de Domínio, como categoria analítica, focalizou-se o conceito de comunidades de prática na condição de unidade de análise instrumental, vista como uma parceria na produção de conhecimento social e com fronteiras dinâmicas. Também recorreu-se à categorização de Ranganathan para identificar as classes mais abrangentes no domínio escolhido e as abordagens *top down* e *bottom-up* para a coleta de termos na construção de Sistemas de Organização do Conhecimento. A proposta de estrutura conceitual resultante desta estratégia é apresentada por meio de um mapa conceitual e em níveis.

Palavras-chave: Organização. Representação. Conhecimento. Estrutura Conceitual. Interdisciplinaridade.

Abstract: The general objective of this paper is to present the identification of strategy of the notional system for psychoactive substances studies in order to develop a Conceptual Structure proposal for an interdisciplinary field, like this one. The strategy adopted refers to the tools and methodologies for building knowledge Organization Systems and conceptual modeling, in particular those related to the development of thesauri. In this strategy, besides on Domain Analysis, as an analytical category, it focused on the concept of communities of practice as the unit of instrumental analysis, seen as a partnership in the production of social knowledge with dynamics borders. Also used the method of categorization established by Ranganathan to identify broader classes in the chosen area and the approaches top down/bottom-up to collect terms for the construction of Knowledge Organization Systems. The Conceptual Structure proposal resulting from this strategy is presented through a conceptual map and at levels.

Keywords: Organization. Representation. Knowledge. Conceptual Structure. Interdisciplinarity.

1 INTRODUÇÃO

No debate sobre o processo de organização do conhecimento, o desafio que é colocado aos profissionais da informação consiste em organizar e representar o conhecimento que não se encaixa na matriz disciplinar.

Reconhecendo as dificuldades na delimitação dos campos do conhecimento em contextos inter, multi e transdisciplinar, defrontamos com a questão da identificação do sistema nocional para a representação e organização do conhecimento.

A ausência do sistema de noções ocasiona a falta de ou a compreensão incorreta das possibilidades de relacionamento entre os termos, de modo que, no caso da construção de Sistemas de Organização do Conhecimento (SOCs), o sistema nocional consiste em parâmetro básico para sua organização e sustentação. Esta ausência compromete a indexação e a economia da atividade documentária, fragmentando-a com questões relativas ao significado e compreensão dos termos, conforme indica a literatura.

Como então construir um sistema nocional na perspectiva interdisciplinar?

O objetivo deste trabalho é a construção de uma Estrutura Conceitual para o domínio identificado como o dos estudos de natureza interdisciplinar das substâncias psicoativas. A construção do sistema nocional aqui proposta adota um paradigma da área da Ciência da Informação, a Análise de Domínio, no qual a informação é algo construído socialmente, por meio da interação do usuário e o seu contexto de informação, em um espaço discursivo constituído pelas principais temáticas de assunto. Também são apropriados aportes de Ranganathan como a definição de categorias e o uso dos modos de formação de assuntos na modelagem conceitual. Para operacionalizar esta análise, a estratégia metodológica foi buscar o conceito de Comunidades de Prática como *locus* da interação usuário e contexto informacional além de ser a unidade instrumental para fazer o recorte da produção interdisciplinar dos estudos de substâncias psicoativas e identificar a garantia literária.

Na seção 2 deste artigo é apresentada a problemática da organização e representação do conhecimento interdisciplinar e a busca de um paradigma no campo da Ciência da Informação, identificado na Análise de Domínio, proposta por Hjørland. O contexto a ser representado, no caso os estudos de substâncias psicoativas, é apresentado na seção 3 onde é resgatada a trajetória histórica do conceito de substância psicoativa a partir dos tipos de uso.

O conceito de Comunidade de Prática e sua referência empírica para atingir os fins propostos neste artigo encontram-se na Seção 4. Nela são detalhados os dois grupos de pesquisa sobre estudos envolvendo substâncias psicoativas e o enfoque de cada um.

Finalmente, na Seção 5 é detalhado o processo de construção da proposta de uma Estrutura Conceitual, lançando mão da definição de categorias e dos relacionamentos e o uso dos modos de formação de assuntos na modelagem conceitual. Também nesta seção é apresentada a proposta de Estrutura Conceitual em 08 facetas e três níveis.

2 O DESAFIO DE REPRESENTAR E ORGANIZAR NO CONTEXTO INTERDISCIPLINAR

Conforme reflexão sobre os desafios para a área de Organização do Conhecimento no século 21, Lara (2013) coloca que

as estruturas e categorias tradicionais não mais dão conta da complexidade contemporânea, o que leva a uma parafernália de conceitos que tentam dar nome e enquadramento àquilo que escapa à categorização. Na tentativa de explicar as ciências, as artes, as técnicas recorre-se a conceitos como interdisciplinaridade, transdisciplinaridade e outros, mas à medida que são, ainda, disciplinares, acabam por reiterar as formas tradicionais de estabelecer as relações entre os fenômenos (LARA, 2013, p.240-241).

A autora pontua ainda “a dificuldade de delimitar os campos do conhecimento, uma vez que constatamos que seus limites estão ‘borrados’ evidencia que “admitimos que eles se entrelaçam, que têm inúmeras interconexões, porém carecemos de ferramentas para enfrentar tais configurações” (LARA, 2013, p.241).

Defrontamos ainda com outro desafio que é a identificação do sistema nocional para a organização e representação do conhecimento. Segundo Cintra et al (2002),

todo e qualquer campo do conhecimento corresponde a um conjunto de noções, sendo que as áreas especializadas têm seu universo nocional devidamente identificado, sempre a partir de um dado ponto de vista, para que seja possível organizá-lo de forma sistemática e inter-relacionada. (CINTRA et al, 2002, p.24).

A ausência do sistema de noções ocasiona a falta de ou a compreensão incorreta das possibilidades de relacionamento entre os termos, de modo que no caso da construção de SOCs, o sistema nocional constitui-se em parâmetro básico para sua organização e sustentação. Esta ausência compromete a indexação e a economia da atividade documentária, fragmentando-a com questões relativas ao significado e compreensão dos termos.

A fragmentação pós-moderna não é disciplinar e sim temática afirma Santos (1988, apud LARA, 2013). Nesta perspectiva “Os arranjos passam a ser mais flexíveis, observando a linguagem das comunidades interpretativas (expressas nos discursos escritos ou orais) e o recurso às metodologias da Terminologia passam a considerar os termos nos contextos de uso”, pontua Lara (2013, p.243).

Como então construir um sistema nocional na perspectiva interdisciplinar?

A construção de um sistema nocional para os estudos de substâncias psicoativas, de natureza interdisciplinar, nos levou a considerar quais os caminhos para o seu estabelecimento no caso de um campo em que se evidencia a seguinte condição:

interação existente entre duas ou mais disciplinas e que pode ir da simples comunicação de ideias até a integração mútua dos conceitos dirigentes, da epistemologia, da terminologia, da metodologia, dos procedimentos, dos dados e da organização da pesquisa e do ensino a eles relacionados (CUNHA; CAVALCANTI 2008, p.210).

Segundo Campos e Moraes (2013, p.73), “a Análise de Domínio fornece apoio metodológico ao mapeamento temático, pois identifica o domínio como uma comunidade discursiva com especial atenção ao contexto onde as trocas informacionais acontecem”. Hjørland e Albrechtsen (1995) identificaram a Análise de Domínio como um paradigma na área da Ciência da Informação, afirmando que: “[...] a melhor forma de entender a informação é através do estudo de domínios do conhecimento através das comunidades discursivas, que são partes da divisão de trabalho na sociedade” (HJORLAND; ALBRECHTSEN, 1995, p.400).

Ainda conforme Campos e Moraes (2013, p.74) a perspectiva da Análise de Domínio percebe a informação como algo construído socialmente, através da interação do usuário com o seu específico contexto de informação, em um espaço discursivo formado pelas principais temáticas de assunto, em seus menores ou mais detalhados ambientes de uso.

Nesta perspectiva, a construção do sistema nocional que nos propusemos terá como correlato a noção de contexto, que segundo Van Dijk (2001, apud CAMPOS; MORAES, 2013, p.75) é construído subjetivamente pelo participante, através de suas interpretações, como experiência única, com linguagem e discurso também únicos. Cada pessoa tem um modelo de contexto com base nas influências e interações com o espaço social, político e cultural, ou também conhecido como, experiências cotidianas. Para a pesquisa em Análise de Domínio, o contexto fornece informações sobre as relações semânticas entre os termos, a palavra e seus múltiplos sentidos, as quais estão imersas em práticas sociais e culturais. O entendimento de um domínio acontece através do estudo da comunidade discursiva, definida por Swales (2009, p.23) considera comunidade discursiva como: “[...] relacionada à produção de textos como uma atividade social que se realiza de acordo com convenções discursivas específicas e revela o comportamento social e o conhecimento dos membros do grupo.”

Para desenvolver a estratégia adotada visando à elaboração do sistema nocional dos estudos de substâncias psicoativas, além da Análise de Domínio, como categoria analítica, focaliza-se o conceito de comunidades de prática como unidade instrumental, resgatando a sua concepção em Wenger (2009) de que do ponto de vista analítico, trata-se do sistema mais simples de produção social do conhecimento social, este entendido como prática. Uma comunidade de prática em si pode ser vista como um sistema social simples, acrescentando ainda que tenha as características de um sistema de aprendizagem social. A aprendizagem como a produção de prática cria limites, porque a partilha de uma história de aprendizagem acaba distinguindo aqueles que estavam envolvidos com aqueles que não estavam. No

entanto, estas fronteiras são dinâmicas e imprevisíveis na concepção desenvolvida por Wenger.

Também recorre-se ao método de categorização ou facetação estabelecido por Ranganathan para identificar as classes mais abrangentes no domínio escolhido e as abordagens *top down* e *botton up* identificadas por Barité (2010, p.130) para a coleta de termos na construção de SOCs.

Segundo Farradane a abordagem denominada *botton up*

[...] parte da coleta e da análise detalhada de casos, objetos, ou expressões da linguagem especializada, para obter uma estrutura organizada de conceitos. Esta abordagem centra-se na coleta de termos que aparecem na documentação ou [...] usados na comunicação e na prática cotidiana de uma comunidade de usuários com aval de especialistas. A coleta de termos pode ser exaustiva ou seletiva considerando os objetivos almejados. Uma vez cumprida esta coleta, os termos são agrupados segundo as classes a que pertencem (FARRADANE, 1950 apud BARITÉ, 2010, p.130).

Já a abordagem *top-down* consiste em, a partir de classes mais gerais, chegar a subclasses cujos termos têm significação mais específica, onde a soma de suas características é maior do que a das classes mais gerais (maior intensão e menor extensão do conceito). Nessa abordagem,

os classificadores seguem um processo de divisão lógica do conhecimento, desde os conceitos mais gerais (normalmente as disciplinas ou subdisciplinas) aos mais particulares (os tópicos específicos). [...] Os sistemas organizacionais do conhecimento assim desenvolvidos são hierárquicos [...] pois cada termo está subordinado a outro, e cada subclasse é um conjunto da classe imediatamente anterior (CANN, 1997 apud BARITÉ, 2010, p.130).

Segundo Barité (2010, p.130), a aplicação dessas duas abordagens concomitantemente para seleção de descritores é fundamental para “uma correta delimitação e organização de um domínio do saber” e é identificada como uma abordagem híbrida. As duas abordagens derivam de dois princípios: o da Garantia Literária e o do Endosso/Garantia do Usuário.

3 O DOMÍNIO DOS ESTUDOS DE SUBSTÂNCIAS PSICOATIVAS

A escolha do domínio identificado como estudos de substâncias psicoativas se deu, conforme Souza Júnior (2013) por tratar-se de uma questão de interesse de saúde pública e social e por ser um domínio em que a interdisciplinaridade se evidencia. Uma estrutura conceitual desse domínio pode facilitar o acesso a informações ligadas à saúde numa perspectiva bem mais abrangente. Trata-se de um domínio interdisciplinar, pois além dos estudos arqueológicos podemos citar o estudo do uso de psicoativos por diversas outras

disciplinas como: Medicina, Psicologia, Veterinária, Direito, Antropologia, Sociologia, Religião, Economia, História, Política, Ética e Farmácia.

Segundo Reis (2007 apud LIMA, 2009, p.39), na Grécia e na Roma antigas existia a noção de *Pharmakon*, quando as bebidas alcoólicas ainda apresentavam ambivalência e eram tratadas como remédio e veneno. Em sua dissertação de Mestrado, Lima (2009, p.41) traça um estudo sobre o consumo dessas substâncias através de uma “visão social de mundo”, na qual as drogas foram sendo, pouco a pouco, constituídas como mercadoria em meio às disputas imperialistas entre grandes potências e frações da classe dominante”. Para essa autora, o tema drogas somente aparece como problema quando existem condições históricas para o desenvolvimento das relações sociais e quando as drogas são tratadas como mercadoria. O uso de drogas para fins de integração cultural não era visto como problema (LIMA, 2009, p.42).

Os debates sobre o uso dessas substâncias são contínuos e muitos governos têm criado restrições sobre a venda dessas substâncias com o intuito de diminuir o abuso no consumo das mesmas.

3.1 De *Pharmacon* à mercadoria

Para entender trajetória do conceito “substâncias psicoativas”, até os dias atuais, retomamos ao *Pharmacon*, que encerra uma noção sem viés proibicionista até a abordagem atual que condena o uso de certas drogas psicoativas e buscamos no trabalho de Lima (2009) o percurso destas substâncias em relação ao seu uso no decorrer da história.

O uso de drogas para fins de integração cultural não era visto como problema (LIMA, 2009, p.42) o café e o ópio já eram produtos em circulação nas rotas comerciais nos séculos XV e XVI, e conforme esta autora, para se compreender como se deu a passagem das plantas para o status drogas torna-se necessário o estudo de como se deu a produção mercantil simples à produção capitalista das drogas como “mercadoria” diferenciando o que é chamado de “valor de uso” e “valor de troca”.

Enquanto a droga está apenas em estágio de “valor de uso”, não atingiu ainda o estágio de “mercadoria”. Para se analisar a condição da *cocaína como mercadoria* é necessário entender os contextos históricos que envolveram a formação das nações americanas através do modelo de colonização espanhola.

A folha de coca como elemento motivador do trabalho dos nativos nas minas de metais preciosos forçou a posição da Igreja quanto à popularização do uso dessa planta entre

os índios (KOUTOUZIS; PEREZ, 1997, p.33, tradução da autora, apud LIMA, 2009, p.65). A folha de coca era aceita como pagamento do dízimo pela Igreja. Assim “a ‘planta mágica’ se tornou um objeto relevante na recém criada classe de comerciantes que se forjava na América” (LIMA, 2009, p.66).

Ainda segundo Lima,

A intervenção da química orgânica sobre esse arbusto, no entanto, permitiu extrair um resíduo oleoso das folhas de coca e depois torná-lo cristalino, alterando sua forma e potencializando os efeitos da excitação com a cocaína refinada. Portanto, os termos coca, folha e droga, respectivamente, referem-se a períodos, a espaços, a usos, a efeitos e, finalmente, a formas mercantis distintas (LIMA, 2009, p.54).

Efetivamente, somente a oligarquia tinha o privilégio do consumo de coca, concedendo-a como recompensa a soldados, mensageiros e camponeses. Mascar sem autorização era considerado crime no Império Inca; sendo assim, tem-se o primeiro caso na história de controle de consumo de substâncias psicoativas em determinadas castas de produção (LEITE; ANDRADE, 1999, p.17 apud LIMA, 2009, p.63).

A folha de coca chegou ao status de mercadoria nos séculos XVIII e XIX em nível de produção mercantil simples. Tornou-se uma mercadoria, na produção mercantil capitalista, no século XX, quando foram criadas as condições nacionais e econômicas propícias e favoráveis ao capitalismo monopolista na América, através da indústria farmacêutica e como mercadoria do tráfico ilícito (LIMA, 2009, p.71).

Lima mostra que “[...] a possibilidade de passagem dessas substâncias de seu estado ‘in natura’ para a condição de drogas, de ‘mercadoria’, requereu, seguramente, o desenvolvimento do comércio, da química e da farmacologia” (LIMA, 2009, p.44).

A síntese da cocaína foi obtida pelo cientista alemão e químico Albert Niemann no ano de 1860.

Também a papoula, ou planta da alegria, bem como o ópio e seus derivados, somente atingiram a condição de mercadoria a partir do momento em que começaram a se desenvolverem condições político-econômicas para permitir sua produção. Esse fato é importante para a análise dos significados sócio-históricos dessas substâncias sob o conceito de mercadoria. Trata-se da análise das relações geoeconômicas e geopolíticas entre Oriente e Ocidente, da ordem feudal à capitalista, e das disputas imperialistas (LIMA, 2009, p.71).

Sinais do uso da papoula foram encontrados em sítios arqueológicos no Centro-Oeste da Europa, desde a era neolítica indicando que é antiga na história humana a procura de substâncias que pudessem alterar o estado de ânimo, e que desde o início da existência humana é comum o uso de plantas especiais. O primeiro registro da papoula foi encontrado a

cerca de 3000 a.C., em um pedaço de argila sob forma de cunha, escrita típica da Mesopotâmia. Esse registro se referia a colheita matinal da “Gil”, ou planta da “alegria”, fazendo referência à sensação de gozo e de prazer (KOUTOUZIS; PEREZ, 1997 apud LIMA, 2009, p.72).

Depois do fim da Inquisição o ópio passa a ser o fármaco predileto das casas reais europeias (ESCOHOTATO, 1997, p.56 apud LIMA, 2009, p.76).

“A partir do século VII, o desenvolvimento do mundo árabe permitiu que o conhecimento medicinal do ópio fosse conhecido na Ásia Central, no sul da Índia e na China em meio às trocas comerciais” (LIMA, 2009, p.78).

Segundo o Grande Livro da Cannabis, de autoria de Rowan (1999, p.5-11), além do uso tradicional no ramo das Ciências medicinais, as fibras do cânhamo possuem diversas outras utilidades. Essas podem ser usadas para a fabricação de tecidos, de papel, na construção de moradias, produção de óleo derivado de suas sementes.

Quanto ao meio ambiente, o cânhamo apresenta as qualidades necessárias para atender a um desenvolvimento sustentável, sem poluir o solo, o ar ou a água e pode atuar como um recurso energético. “É um produto auto-suficiente e local, sem depender de países estrangeiros para explorá-la e o seu processo de produção poderia gerar empregos (ROWAN, 1999, p.13).

Na área da saúde, as pesquisas médicas que apóiam o uso da *cannabis sativa* em tratamentos terapêuticos são volumosas. Servem para o tratamento do glaucoma, como antiemético, para tratar dificuldades respiratórias, como antibiótico, antiartrítico, antidepressivo, para controlar inflamações, como analgésico, para tratar o alcoolismo, para alívio da insônia, alívio do herpes, da enxaqueca e para o uso ginecológico na doença de *hyperemesis gravidarum*, que é uma forma de enjôo matinal que a mulher grávida sofre de constantes náuseas e vômitos (ROWAN, 1999, p.33-37).

O uso de substâncias psicoativas pode ser de diferentes propósitos. As diferenças culturais são um forte fator que determinam o uso dessas substâncias. Por exemplo: o uso do *crack* se dá entre indivíduos de baixo poder aquisitivo, provenientes de comunidades nas quais o índice de analfabetismo é muito grande.

Algumas dessas substâncias são de uso controlado ou ilegal, enquanto outras podem ser usadas para fins religiosos ou terapêuticos. O consumo de bebidas alcoólicas e o uso de substâncias para dormir são exemplos de consumo social. A cafeína é a substância psicoativa mais consumida no mundo e seu uso é legalizado. O Brasil é o segundo maior produtor de café no mundo e 85% dos brasileiros consomem café no desjejum. O uso dessas substâncias,

além dos fins medicinais pode ser para fins recreativos (alteração proposital da consciência), religiosos (fazer contato com entidades) e científicos (WIKIPÉDIA, 2013).

Essas substâncias também podem ser usadas para fins religiosos e medicinais. Como exemplo do uso religioso das substâncias psicoativas pode-se citar o ritual de cura do Santo Daime (ayahuasca), o uso da *Lophophora williamsii* ou mescalina (peote/ erva do diabo) e cogumelos.

4 COMUNIDADES DE PRÁTICA EM ESTUDOS INTERDISCIPLINAR DE SUBSTÂNCIAS PSICOATIVAS

Uma comunidade de prática pode ser vista como um sistema de aprendizagem social e exhibe muitas características dos sistemas de forma mais geral: estrutura emergente, relacionamentos complexos, auto-organização, limites dinâmicos, negociação em curso de identidade e cultural, para mencionar alguns. Num sentido, é a mais simples unidade social que tem as características de um sistema de aprendizagem social. (WENGER, 2009).

Ainda na perspectiva desenvolvida por Wenger (2009), Comunidades de prática, naturalmente, não estão isoladas; elas são parte de sistemas sociais mais amplos que envolvem outras comunidades (bem como outras estruturas tais como projetos, instituições, movimentos, ou associações). Assim, o mundo social inclui práticas inumeráveis; e vivemos e aprendemos através de uma multiplicidade de práticas

Como campo empírico de referência dos estudos interdisciplinares sobre estas substâncias, encontramos duas comunidades de produção de conhecimento sobre estudos de substâncias psicoativas ligadas às universidades brasileiras que se encaixam no conceito de comunidade de prática de Wenger.

Uma delas é o Núcleo de Estudos Interdisciplinares sobre Psicoativos (NEIP) que desenvolve pesquisa sobre substâncias psicoativas reunindo estudiosos das áreas de Ciências Humanas, vinculados a diversas instituições, para promover uma reflexão contínua sobre o tema.

O site do NEIP⁴⁰ funciona como um espaço de diálogo acadêmico útil entre pessoas que desejam debater sobre o tema. Esse Núcleo adota uma postura antiproibicionista quanto às regularizações governamentais e é filiado à Associação Brasileira de Estudos Sociais de Substâncias Psicoativas (ABESUP).

⁴⁰ <http://www.neip.info/index.php/content/view/1.html>

Ainda conforma o site, o NEIP foi criado em 2001 e defende que essa área tem um papel fundamental no debate sobre o “problema das drogas” dominado pelas ciências da saúde e por posicionamentos muitas vezes preconceituosos.

A abordagem teórica adotada por esse grupo de pesquisa é marcada pela interdisciplinaridade entre saberes como o filosófico, o sociológico, o antropológico, o econômico, o artístico e histórico e pelo diálogo entre esses diversos campos do saber.

O NEIP defende a ideia de que qualquer indivíduo possui o direito de experimentar sua própria condição de ser da maneira que quiser, sem sofrer a interferência de controle externo, político e totalizante. A proibição e a destruição desses hábitos milenares por parte da sociedade de consumo sustentam as políticas do Núcleo, que assume uma postura ética de crítica ao proibicionismo e defesa das experiências individuais e coletivas com psicoativos, entendidas como atitudes humanas e sociais legítimas. Para isso problematiza o tema da ilegalidade, que considera um pressuposto para o combate a grupos sociais, hábitos individuais e práticas seculares.

Também em pesquisa na Internet, encontramos o site do Grupo Interdisciplinar de Estudos sobre Substâncias Psicoativas (GIESP) que é formado por alunos e professores das diferentes áreas humanas e de saúde que estudam o fenômeno do uso de substâncias psicoativas. O grupo existe desde 2005 e está inserido no Diretório de Grupos do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) certificado pela Universidade Federal da Bahia (UFBA) e liderado pelo pesquisador Edward MacRae. As temáticas de interesse do Grupo são: Educação e Prevenção ao Uso; Redução de Danos e Práticas Terapêuticas; Legislação e Políticas Públicas; Ética em Pesquisa e na Atenção aos Consumidores; Usos Tradicionais e Espirituais; Produção, Comercialização e Consumo.

O GIESP também se propõem a discutir o assunto em abordagem interdisciplinar, possui produção sobre o tema e defende uma abordagem biopsicosocial quanto aos seus estudos.

5 A CONSTRUÇÃO DA ESTRUTURA CONCEITUAL

Foi feito o mapeamento da literatura sobre substâncias psicoativas produzida por estes dois grupos que nos forneceu o campo semântico destes estudos. Tal mapeamento possibilitou levantar as principais categorias e subcategorias de análise que deveriam estar presentes na matriz a ser aplicada na análise documentária do campo empírico proposto para fins de elaboração de uma proposta de estrutura conceitual.

No processo de categorização buscou-se compreender a dimensão complexa deste domínio e foram encontradas na literatura algumas pistas para a escolha de categorias. No entanto, além do mapeamento da literatura, nos baseamos no método de facetação e para alguns termos nos baseamos nos princípios para a formação de assuntos propostos em Ranganathan.

Conforme Piedade (1983, p.25), “a palavra faceta às vezes é empregada no sentido aristotélico de categoria para designar as categorias fundamentais, as classes gerais de fenômenos ou grandes grupos que podem ser constatados no exame de um assunto”.

Ranganathan (1967, p.88) define Faceta como “um termo genérico usado para denotar algum componente –pode ser um assunto básico ou um isolado- de um assunto composto, tendo, ainda a função de formar renques, cadeias”. Segundo Campos e Gomes (2003), Ranganathan propõe que os conceitos sejam organizados em renques e cadeias subordinados a uma categoria. Renques e Cadeias são denominações dadas por Ranganathan para diferenciar séries verticais e horizontais dos conceitos e revela a estrutura de organização por meio das relações de gênero/espécie e todo/parte.

Os princípios para a formação de assuntos estabelecidos por Ranganathan são úteis quando não existem relações hierárquicas de gênero/espécie e todo/parte. São considerados de interesse para os estudos da interdisciplinaridade:

Dissecação é cortar um universo de entidades em partes. [...] Laminação é a construção por superposição de facetas sobre facetas. [...] Desnudação é a diminuição progressiva da extensão e o aumento da intensão de um assunto básico ou uma ideia isolada. [...] Reunião Livre é a reunião de dois ou mais: a) assuntos básicos ou compostos, b) ideias soladas.[...] E Superposição liga duas ou mais ideias isoladas que pertencem ao mesmo universo de idéias isoladas [...] (RANGANATHAN, 1967 apud SILVA, 2007, p.22-23).

Segundo Kumar (1981apud SILVA, 2007, p.21), os estudos de modos de formação e estrutura de assuntos também auxiliam na organização e sistematização do sistema nocional, pois permite prever e entender os vários tipos de inter-relações entre os assuntos. Ainda segundo Kumar, outros autores seguiram a linha de Ranganathan e ampliaram o estudo dos modos de formação de assuntos [...] São eles:

Aglomerção ou Agrupamento [...] é formado através de diferentes assuntos básicos ou assuntos compostos.

Destilação é uma disciplina pura e é envolvida a um assunto principal para apresentar diversos assuntos compostos”.

Divisão ou Fissão é um assunto básico ou isolado subdivididos em um processo considerado [de] fragmentação.

Fusão: “é caracterizada quando dois ou mais assuntos principais são agrupados de tal maneira que possam compor um novo assunto principal.

Permite a formação de assuntos multidisciplinares” (KUMAR, 1981, apud SILVA, 2007, p.44).

Através dos modos de formação de assuntos propostos, foi possível classificar alguns descritores que compõem a estrutura conceitual proposta, conforme abaixo:

QUADRO 1 – Organização de descritores com base nos modos de formação de assuntos

Modo de formação de assunto	Descritores dos Estudos sociais do uso de substâncias psicoativas
Laminação (Construção por superposição de faceta sobre faceta)	Ex: .Efeito . Efeito Farmacológico ...Dano Farmacológico ... Ginecomastia (aumento das glândulas mamárias masculinas por uso de medicamento com hormônio)
Dissecação (cortar um universo de entidades em partes que tenham posição coordenada, como na formação de um renque)	Ex: .Controle .. Autocontrole .. Heterocontrole .. Sociocontrole

Fonte: a autora

Gopinath & Seetharama (1979, apud SILVA, 2007, p.26) apontam para outra estratégia na elaboração do *design* de um sistema de classificação a partir do estudo do desenvolvimento histórico de uma área de conhecimento reconhecendo as variações dos assuntos (ou conceitos para o nosso caso) em cada fase do seu desenvolvimento.

Com base no quadro abaixo, que representa a matriz de formação de assuntos numa perspectiva histórica, podemos observar que a primeira relação formada é a de Reunião Livre, que ocorreu na Grécia e Roma antiga, quando o assunto substâncias psicoativas ainda estava em fase inicial de formação com a noção de *pharmacon*.

A segunda relação é a de Laminação, quando a visão social de mundo introduzida pela Economia política de Marx, atrela o significado do conceito de drogas ao conceito de mercadoria.

A terceira relação é a de Fissão, quando são traçadas fronteiras mais rígidas entre os estudos do uso de substâncias psicoativas. Surge o conceito de “droga”.

A quarta relação demonstrada no quadro matriz de formação de assuntos é a de Fusão. Essa relação mostra a fusão de três assuntos principais que dão origem a um novo assunto básico: os aspectos biopsicosociais dos estudos de substâncias psicoativas.

QUADRO 2 – Organização de descritores com base nos modos de formação de assuntos em uma perspectiva histórica

Modo de formação de assunto	Estudos sociais do uso de substâncias psicoativas
Reunião Livre (Assunto em estágio inicial)	Na Grécia e Roma antiga havia a noção de <i>pharmacon</i> , quando ainda havia a ambivalência, de forma simultânea, como remédio e veneno (LIMA, 2010, p.39).
Laminação (Construção por superposição de faceta sobre faceta)	Através da visão social de mundo introduzida pela Economia Política Marxista, o conceito de <i>pharmacon</i> é substituído pelo de alcalóide (substância sintetizada), transforma-se em droga, que começa a ser tratada como mercadoria em meio às disputas imperialistas entre grandes potências e frações da classe dominante.
Fissão (Subdivisão de um assunto básico ou isolado)	A partir do final do século XIX, o significado social da droga vai perdendo o seu constitutivo contraditório para assumir fronteiras mais rígidas. Se o consumo da droga lícita for prescrito por médicos ou for uso recreativo, seu consumidor estará seguro. Se for o consumo de uma droga controlada e não prescrita, seu consumidor estará deslocado para o território do veneno/sofrimento (LIMA, 2010, p.41).
Fusão (Indicação de um novo assunto básico através da união entre assuntos principais)	Na modernidade, ocorre uma fusão entre as abordagens biomédicas, psicológicas, sociais e econômicas nos estudos de uso de substâncias psicoativas (abordagem biopsicossocial).

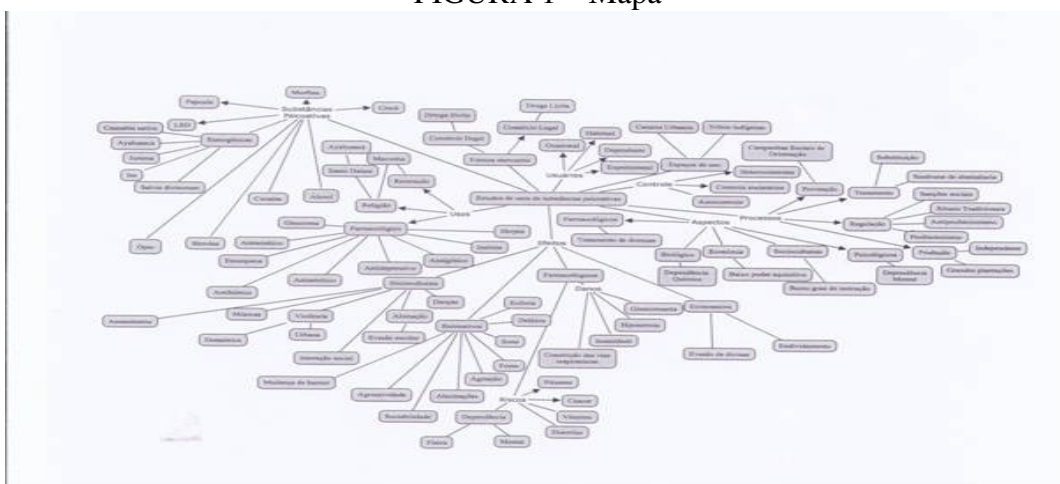
Fonte: SOUZA JR. (2013)

Os termos coletados foram estruturados e registrados em um mapa conceitual que foi produzido com o auxílio do aplicativo *Cmap*. Esse aplicativo foi desenvolvido pelo *Institute for Human and Machine Cognition da UWF* – Universidade de West Florida. Este software utiliza como ferramenta de organização dos conceitos, um diagrama hierárquico, apresentando a informação de forma descendente de importância. A informação mais importante, ou domínio, é colocada no início da cadeia hierárquica. Para elaboração deste mapa foram adotadas divisões utilitárias, que servem somente como elementos organizadores do mesmo.

Alem de estrutura conceitual do domínio de psicoativos, foi elaborado um glossário com a definição dos termos selecionados como descritores. A base documentária analisada foi definida a partir da produção literária do GIESP e NEIP como: Livros, textos, e teses.

A seguir, na figura 1 apresentamos o Mapa Conceitual da proposta e na figura 2 a proposta de Estrutura Conceitual para representação do domínio interdisciplinar de Estudos de Substâncias Psicoativas. (103 Termos) dividida em 08 (oito) categorias principais com 03 (três) níveis hierárquicos.

FIGURA 1 – Mapa



Fonte: Souza Júnior, 2013

FIGURA 2 – Estrutura Conceitual em Níveis

Faceta	Aspecto	Controle	Efeito	Espaço de Uso	Processo	S.Psicoativa	Uso	Usuário
NIVEL 1	1.Econômico	1.Autocontrole	1.Econômico	1.Centro Urbano	1.Forma Mercantil	1.Álcool	1.Farmacológico	1.Experimental
		2.Heterocontrole	2.Farmacológico	2.Tribo Indígena	2.Regulação	2.Cocaina	2.Recreativo	2.Dependente
	2.Sociocultural	3.Sociocontrole	3.Recreativo		3.Prevenção	3.LSD	3.Religioso	3.Ocasional
			4.Sociocultural		4.Terapia	4.Papoula	4.Substância Enteogênica	4.Habitual
Nível 2	Nível 3	1.1 Baixo poder Aquisitivo	1.1Evasão de Divisas		1.1Tráfico		1.1Medicamento	
			1.2Endividamento		1.2Comércio Lícito	2.1 Crack	1.2 Tratamento	
			2.1 Dano Farmacológico		2.1 Antiproibicionismo	4.1 Opio	3.1 Ritual	
			2.2 Risco Farmacológico		2.2 Proibicionismo	5.1 Ayahuasca		
			3.1 Agitação		2.3 Sanção	5.2 Cannabis		
			3.2 Alucinação		3.1 Campanha social de orientação	5.3 Jurema		
			3.3 Delírio		4.1 Substituição	4. Salvia		
			3.4 Euforia		4.2 Síndrome de Abstinência			
			3.5 Fome					
			3.6 Humor					
3.7 Sociabilidade								
3.8 Sono								
4.1 Agressividade								
4.2 Alienação								
4.3 Interação								
4.4 Produção Cultural								
4.5 Violência								
2.1.1 Ginecomastia			1.1.1 Droga Ilícita	5.2.1 THC	1.1.1 Analgésico			
2.1.2 Hipotermia			1.2.1 Droga Lícita		1.1.2 Antiarrítmico			
2.1.3 Insanidade					1.1.3 Antibiótico			
2.1.4 Constrição					1.1.4 Antidepressivo			
2.2.1 Cancer					1.1.5 Antimético			
2.2.2 Dependência					2.1.1 Enxaqueca			
2.2.3 Diarréia					2.2.2 Glaucoma			
2.2.4 Intoxicação					1.2.3 Herpes			
2.2.5 Náusea					2.2.4 Inanônia			
2.2.6 Vômito					3.1.1 Santo Daime			
4.4.1 Música								
4.4.2 Dança								
5.5.1 Assassinato								
5.5.2 Violência Doméstica								
5.5.3 Violência Simbólica								
5.5.4 Violência Urbana								

Fonte: a autora

6 CONCLUSÕES

A impossibilidade humana de acumulação e processamento de todo o conhecimento disponível reforça a nossa dependência cada vez maior da eficaz organização e representação do conhecimento produzido. A abordagem participativa na organização do conhecimento por meio de comunidades de prática, discursivas ou virtuais é uma tendência em sociedades que se caracterizam pela complexidade e fragmentação do conhecimento.

A adoção do conceito de Comunidades de Prática como unidade instrumental para operacionalizar a nossa proposta se mostrou compatível com a necessidade de representação de um sistema de produção social de conhecimento cujas fronteiras se caracterizam pelo entrelaçamento em múltiplas conexões e é preciso adotar este novo paradigma, da Análise do Discurso, para se aproximar ao máximo do campo semântico dos novos conhecimentos em produção, visando a sua representação e organização. O trabalho desenvolvido mostrou também que mesclar novas estratégias aos aportes tradicionais da área, como no caso de Ranganathan, resulta em um ganho de escala. A Estrutura Conceitual aqui apresentada é uma primeira proposta que deve se desdobrar em um estudo mais apurado para a construção de um tesouro para estes pesquisadores.

REFERÊNCIAS

- BARITÉ, M. *et al.* Garantia Literária: elementos para uma revisão crítica após um século. **Transinformação**, Campinas, v. 22, n. 2, p. 123-138, maio/ago. 2010
- BIASI-RODRIGUES, B.; ARAÚJO, J. C.; SOUSA, S. C. T. 2009. Análise de gêneros na abordagem de Swales: princípios teóricos e metodológicos. In _____. (Org.). **Gêneros textuais e comunidades discursivas: um diálogo com John Swales**. Belo Horizonte: Autêntica, 17-32
- CAMPOS, M.L. A; GOMES, H.E. Organização de domínios de conhecimento e os princípios ranganathanianos. **Perspectiva em Ciência da Informação**. Belo Horizonte, v. 8, n. 2, p.150-163jul/dez. 2003
- CAMPOS, M.L. A; MORAES, R.P.T.de. Mapeamento conceitual sob o enfoque da análise de domínio: uma discussão de conceitos fundamentais. In: DODEBEI, V. GUIMARÃES, J.A.C. (Orgs.). **Estudos Avançados em Organização do Conhecimento**, v. 2, Rio de Janeiro: Isko-Brasil p.73-78, 2013.
- CAMPOS, M. L.A. Linguagem documentária: teorias que fundamentam sua elaboração. 2. ed. Niterói, RJ: EdUFF, 2002.
- CANN, J. Principles of classification: suggestions for a procedure to be used by ICIS. In: **Developing international classification tables for the construction industry**. 1997. Disponível: <<http://www.icis.org/siteadmin/rtdocs/images/5.pdf>>.
- CINTRA, A. M. M. *et.al.* **Para entender as linguagens documentárias**. 2. ed. rev. e ampl.; São Paulo: Polis, 2002.

- CUNHA, M. B. da; CAVALCANTI, C. R. O. **Dicionário de biblioteconomia e arquivologia**. Brasília, DF: Briquet de Lemos/ Livros, 2008.
- ESCOHOTADO, A. O livro das drogas: usos e abusos, preconceitos e desafios. SP, Dynamis Editorial, 1997.
- FARRADANE, J.E.L. A scientific theory of classification and indexing. **Journal of Documentation**, v. 6, p.83-99, 1950.
- GOPINATH, M. A.; SEETHARAMA, S. Interdisciplinary subjects and their classification. In: Neelameghan, A. (Ed.). **Ordering systems for global information networks**. Bangalore: FID/CR Committee and Sarada Ranganathan Endowment for Library Science, p. 121-134, 1979.
- HJØRLAND, B.; ALBRECHTSEN, H. Toward a new horizon in Information Science: domain-analysis. **Journal of the American Society for Information Science**, v.46, n.6, p.400-425, Jul. 1995.
- KOUTOUZIS, M. & PEREZ, P. **Atlas Mondial des Drogues**. Paris, Presses Universitaires de France, 1997.
- KUMAR, K. **Theory of classification**. 2.ed. Delhi: Vikas Publishing House, 1981.
- LARA, M.G. G. Problemas da organização do conhecimento na contemporaneidade. In: In: DODEBEI, V. GUIMARÃES, J.A. C. (Orgs.). **Estudos Avançados em Organização do Conhecimento**, v. 2, Rio de Janeiro:Isko-Brasil, p.238-245, 2013.
- LEITE, M. C.; ANDRADE, A. **Guerra de Cocaína e crack: dos fundamentos ao tratamento**. Porto Alegre, Ed. Artes Médicas Sul Ltda., 1999.
- LIMA, R.C. **Uma história das drogas e do seu proibicionismo transnacional: relações Brasil-Estados Unidos e os organismos internacionais**. 2009.365.f.Tese (Doutorado em Serviço Social)-Universidade Federal do Rio de Janeiro, Rio de Janeiro,2009. Disponível em: <<http://pt.scribd.com/doc/36564062/Uma-historia-das-drogas-e-do-proibicionismo-transnacional-Tese-Lima-2009-UFRJ>>. Acesso em: 26 out./2013.
- PIEIDADE, M. A. R. **Introdução à teoria da classificação**. 2. ed. rev. e ampl. Rio de Janeiro: Interciência, 1983.
- RANGANATHAN, S. R. **Prolegomena to library classification**. Bombay: Asia Publishing House, 1967.
- REIS, T.R. Fazer em grupo o que eu não posso fazer sozinho: indivíduo, grupo e identidade social em Alcoólicos Anônimos. Tese (Doutorado em Serviço Social) UFRJ. Rio de Janeiro, 2007.
- ROWAN, R. **O grande livro da Cannabis: guia completo de seu uso industrial, medicinal e ambiental**. Rio de Janeiro: Jorge Zahar, 1999.
- SANTOS, B. S. Um discurso sobre as ciências na transição para uma ciência pós-moderna. **Estudos avançados**, São Paulo, v.2, n.3, maio/ago. 1988. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-40141988000200007 . Acesso em: 3 abr. 2013.

SILVA, R.P. **A interdisciplinaridade e os aspectos conceituais e de representação**: análise da área multidisciplinar da tabela de áreas do conhecimento em uso pela CAPES. Dissertação (mestrado em ciência da informação). IBICT-UFF. Rio de Janeiro, 2007.

SILVA, R.P. **A interdisciplinaridade e os aspectos conceituais e de representação**: análise da área multidisciplinar da tabela de áreas do conhecimento em uso pela CAPES. Dissertação (mestrado em ciência da informação). IBICT-UFF. Rio de Janeiro, 2007.

SOUZA JR, E. **Proposta de Estrutura Conceitual para um domínio interdisciplinar**: o caso dos estudos de substâncias psicoativas. 2013.99.f. Trabalho de Conclusão de Curso (Bacharelado)-Escola de Biblioteconomia, Universidade Federal do Estado do Rio de Janeiro, Rio de Janeiro, 2013

SWALES, J. M. Sobre modelos de análise do discurso. In: BIASIRODRIGUES, B.; ARAÚJO, J. C.; SOUSA, S. C. T. de. (Orgs.). **Gêneros textuais e comunidades discursivas**: um diálogo com John Swales. Belo Horizonte: Autêntica, p. 33-46, 2009.

VAN DIJK, T. A.. Algunos principios de una teoria del contexto. **ALED, Revista Latinoamericana de estudios del discurso**, v.1(1), p.69-81. 2001.

WENGER, E. Communities of practice and social learning systems: the career of a concept. In: Blackmore. C. (Editor) **Social Learning Systems and communities of practice**. Springer, Verlag and the Open University.2009. Disponível em: <<http://wenger-trayner.com/resources/publications/cops-and-learning-systems/http://wenger-trayner.com/wp-content/uploads/2012/01/09-10-27-CoPs-and-systems-v2.0.pdf>>. Acesso em: 21 dez. 2013.

IDENTIFICAÇÃO DE TIPOLOGIA DOCUMENTAL COMO PARÂMETRO PARA CLASSIFICAÇÃO

IDENTIFICATION DOCUMENT TYPE AS A BENCHMARK FOR CLASSIFICATION

Silvia Lhamas de Mello
Ana Célia Rodrigues

Resumo: Apresenta o tema da classificação de documentos de arquivos fundamentados na metodologia de identificação arquivística com o objetivo de analisar as relações da identificação com a classificação no âmbito dos arquivos universitários, especificamente da área de pessoal da Universidade Federal do Rio de Janeiro. Descreve a base teórico-metodológica da Diplomática e da Tipologia Documental como parâmetro para identificação e classificação na Arquivística e suas relações no contexto da gestão de documentos. Apresenta uma contextualização dos estudos da Diplomática, a transição de seu objeto até o contraponto entre a Tipologia Documental e Diplomática Contemporânea. Busca compreender os fundamentos teóricos e metodológicos da identificação arquivística que amparam a construção de planos de classificação, observando sua aplicabilidade anterior às funções de classificação e avaliação. Apresenta um estudo de caso de aplicação da metodologia da identificação na Divisão de Gestão Documental e da Informação da Universidade federal do Rio de Janeiro. Descreve a Universidade Federal do Rio de Janeiro e seus acervos, apresentando um estudo de identificação de órgão produtor, os elementos orgânico e funcional da Pró-Reitoria de Pessoal da Universidade Federal do Rio de Janeiro com base na estrutura organizacional, no regimento interno e nas entrevistas, com a finalidade de elaborar uma proposta de classificação por amostragem e os procedimentos metodológicos desenvolvidos.

Palavras-chave: Classificação. Diplomática. Identificação Arquivística. Tipologia. Arquivos Universitários.

Abstract: Presents the issue of classification of archival documents anchored in archival identification methodology in order to analyze the relationship of identification with the classification under the files university, specifically in the personal area of Federal University of Rio de Janeiro. Describes the theoretical and methodological bases of the Diplomatic and documentary typology as a parameter for identification and classification Archives and relations in the context of document management. Presents a contextualization of Diplomatic Studies, the transition of its object to the contrast between the types of documents and contemporary diplomatic. Seeks to understand the theoretical and methodological foundations of archival identification that support the construction of classification schemes, noticing its applicability before the functions of classification and evaluation. Presents a case study of application of the methodology of identification in the Division of Document Management and Information Federal University of Rio de Janeiro. Describe the Federal University of Rio de Janeiro and its collections, presenting a study of identification organ producer, organic and functional elements of the field of Personal Federal University of Rio de Janeiro based on organizational structure, bylaws and in the interviews, in order to prepare a proposal for classification sampling and methodological procedures developed.

Keywords: Classification. Diplomatic. Archival Identification. Typology. University Archives.

1 INTRODUÇÃO

O tema em questão, classificação na Arquivística, encontra cada vez mais visibilidade e importância no âmbito das discussões teóricas da área, porque se trata de uma função primordial para o desenvolvimento da gestão documental nas instituições.

É a gestão de documentos, configurada por um conjunto de procedimentos desenvolvidos para controlar os documentos, desde sua produção até sua destinação final (eliminação ou guarda permanente), que contribui para o bom funcionamento dos processos informacionais de uma instituição, uma das funções que sustentam o programa de gestão de documentos é a classificação.

No contexto da gestão de documentos, os estudos dos tipos documentais são imprescindíveis para a realização das funções que sustentam o programa, ou seja, a classificação e a avaliação. A identificação arquivística se define como o “processo de reconhecimento, sistematização e registro de informações sobre arquivos com vistas ao seu controle físico e/ou intelectual” (DICIONÁRIO BRASILEIRO DE TERMINOLOGIA ARQUIVÍSTICA, 2005, p.104). Estes estudos desenvolvidos no momento da identificação permitirão o reconhecimento das estruturas administrativas e das suas atribuições, elementos que ficam registrados nos tipos documentais, cuja metodologia encontra, nos parâmetros da Diplomática Contemporânea, a base teórica para a construção dos instrumentos de gestão.

A análise das competências, funções e atividades que se articulam nos procedimentos administrativos foi incorporada ao método da Diplomática Contemporânea para analisar o documento de arquivo, ou seja, o tipo documental. E sua finalidade é desvendar o vínculo arquivístico, isto é, a proveniência e a organicidade do documento.

No Brasil, a falta de um modelo metodológico consistente para a construção de instrumentos de classificação é uma realidade e a busca por meios que a transforme, tem sido um dos esforços de pesquisa na área Arquivística. A diferença entre princípios e critérios adotados e a consequente diferença entre os resultados obtidos, vem gerando um debate teórico e metodológico na área, sobre a pertinência e consequências da aplicabilidade de instrumentos de classificação pré-estabelecidos nos arquivos.

Desde que trabalho como arquivista, desenvolvo atividades relacionadas à classificação de documentos. O interesse pelo tema foi aguçado quando ingressei como servidora pública na Universidade Federal do Rio de Janeiro (UFRJ), com a tarefa de resolver o problema de uma imensidão de papéis acumulados e de implantar um programa de gestão documental para o arquivo universitário, cujas características são muito particulares, pois está voltado para a divulgação do conhecimento científico. E ainda, como a universidade é parte

integrante do Sistema Nacional de Arquivos (SINAR), tem a obrigatoriedade legal de utilizar a Resolução nº 14 do Conselho Nacional de Arquivos (CONARQ), onde estão dispostos os instrumentos de gestão documental.

Os problemas relacionados com a utilização de instrumentos de classificação pré-definidos na implantação da gestão documental, em uma instituição, são vários, um deles é a falta de consistência entre os assuntos estabelecidos e a realidade encontrada nos documentos. Desta forma, os estudos da tipologia documental realizados no contexto da identificação arquivística, podem oferecer possibilidades teórico-metodológicas de realização de uma efetiva gestão documental, especificamente, fornecendo as bases para a classificação.

Dar acesso aos documentos públicos é uma atividade constitucional, pois segundo a Constituição Federal de 1988, “cabem à administração pública, na forma da lei, a gestão da documentação governamental e as providências para franquear sua consulta a quantos dela necessitem” (artigo 216, §2º). A pesquisa proposta torna-se relevante ao ponto em que uma instituição da importância da UFRJ não pode ferir um ato constitucional. Portanto, para dar acesso ao rico acervo universitário é necessária a existência, na instituição, de propostas que viabilizem a implantação da gestão documental, preservação dos acervos e acesso aos documentos produzidos e acumulados pela universidade. Para tanto, o desenvolvimento da pesquisa se justifica trazendo à tona reflexões teóricas e práticas acerca do fazer arquivístico, possibilitando o melhoramento das atividades desenvolvidas não só pela equipe da Divisão de Gestão Documental e da Informação (DGDI), mas também contribuindo socialmente com aqueles que buscam referências sobre o tema e com as características dos documentos universitários.

A importância da informação é atualmente um fato aceito pelas instituições. A necessidade de tratamento da informação e de tornar acessível um acervo em constante crescimento são problemas que também permeiam a Ciência da Informação. Ao longo de sua história a Ciência da Informação vem desenvolvendo suas práticas no que tange à organização e à recuperação da informação. Não distante desta realidade, a Arquivística se insere neste contexto, uma vez que busca, na relação interdisciplinar com a Ciência da Informação e outras áreas do conhecimento como a Administração, soluções para seus problemas de organização e recuperação das informações orgânicas. Neste cenário, o documento atua como mediador num processo informacional, de transmissão de mensagens para um agir informado onde a sociedade pode apropriar-se através da produção, armazenamento, transferência e recuperação de potencialidades informativas (MARCONDES, 2011, s/p).

Os temas que são abordados nesta pesquisa estão diretamente relacionados com a gestão de documentos arquivísticos, suas construções teóricas e metodologias para aplicação nos programas institucionais, principalmente no que tange à classificação de documentos. O contexto apresentado, busca nas teorias e metodologias da área, tentativas de solucionar as lacunas deixadas até o momento referentes à classificação em arquivos.

Diante deste cenário, as questões que norteiam esta pesquisa são: a metodologia da identificação, realizada na perspectiva da Tipologia Documental, pode fornecer parâmetros para a elaboração de planos de classificação no âmbito dos arquivos universitários? Esses parâmetros podem ajudar na utilização de instrumentos de gestão pré-definidos? Os regimentos são suficientes para identificar o órgão produtor e contextualizar a tipologia documental no plano de classificação? Pretende-se encontrar respostas que facilitem a tarefa da classificação de documentos no contexto universitário.

O objetivo geral desta pesquisa é analisar as relações da identificação com a classificação no âmbito de arquivos universitários, especificamente na área de pessoal da UFRJ. Como desdobramentos deste objetivo geral, são objetivos específicos: apresentar os princípios teóricos e metodológicos da identificação arquivística, realizada nos parâmetros da Tipologia Documental e sua pertinência para a classificação; descrever os princípios teóricos e metodológicos da classificação na Arquivística no contexto da gestão de documentos; aplicar a metodologia da identificação para elaborar um plano de classificação funcional para a área de pessoal da UFRJ, compatibilizando seu uso com o Código de Classificação da Atividade-Meio para a Administração Pública Federal.

2 METODOLOGIA

O desenvolvimento desta pesquisa se caracteriza metodologicamente por um estudo qualitativo exploratório, de natureza documental e aplicada, com o objetivo de reunir informações acerca dos métodos e teorias da identificação e da classificação na Arquivística relacionada com os parâmetros da Tipologia Documental. Para atingir os objetivos propostos, esta pesquisa aliou uma dimensão teórica a uma aplicada. Na dimensão teórica, realizou-se um levantamento bibliográfico e revisão de literatura sobre os temas: Diplomática, Tipologia Documental, Identificação Arquivística, Gestão de Documentos e Classificação na Arquivística para conhecer os fundamentos teórico-metodológicos e embasar o aspecto aplicado desta pesquisa, resultando nos capítulos teóricos da pesquisa.

Na dimensão aplicada, a pesquisa adotou dois procedimentos:

Procedimentos de Identificação

Identificação de órgão produtor:

a) elemento Orgânico: levantamento de dados através da análise da área administrativa nas fontes de pesquisa: organogramas e regimentos;

b) elemento Funcional: levantamento das competências, funções, atividades e tarefas através de fontes: regimentos, estatutos, normas de serviço e entrevistas;

c) aplicação dos procedimentos de estudo de órgão produtor proposto por Ana Célia Rodrigues (2012)⁴¹, desenvolvidos no Programa de Gestão Documental do Estado do Rio de Janeiro (PGD-Rio).

QUADRO 1 – Estudo de identificação de órgão produtor

Competência	Função	Atividade	Tarefa	Tipo Documental
-------------	--------	-----------	--------	-----------------

Identificação de tipologia documental:

Levantamento dos tipos documentais aplicando o formulário de denominação de tipos documentais, elaborado com alguns itens da proposta de Bellotto (2008, p.79), através da análise dos processos referentes ao pessoal que se encontram na DGDI da UFRJ com o objetivo de denominar o tipo documental de maneira adequada e conseqüentemente, a denominação da série documental.

QUADRO 2 - Formulário para denominação do tipo documental

Entidade Produtora	Espécie Documental	Ação (verbo) / Atividade	Denominação do Tipo documental Espécie+verbo+objeto da ação
--------------------	--------------------	-----------------------------	--

✓ Procedimentos de Classificação

- Análise de Dados e Resultados: elaborou-se uma proposta de plano de classificação para a área de pessoal da UFRJ. Esta proposta visou a validação dos estudos de identificação na construção de planos de classificação.

3 DIPLOMÁTICA E TIPOLOGIA DOCUMENTAL

A Diplomática tem sua origem no século XVII, na França, para resolver problemas de disputas de terra entre ordens religiosas. Beneditinos e Bolandistas faziam parte desta disputa levando o frei Jean Mabillon a desenvolver uma metodologia que pudesse comprovar a

⁴¹ RIO DE JANEIRO (Estado). Governo do Estado do Rio de Janeiro. **Manual de Gestão de Documentos**. Rio de Janeiro, 2012. 108 p.

autenticidade dos títulos de terra existentes nos arquivos de sua abadia (GARCIA PIQUERAS apud RODRIGUES, 2008, p.121).

Considera-se a Diplomática como a disciplina dos documentos, onde seu objeto, os documentos diplomáticos, “são os de natureza jurídica que, refletem no ato escrito as relações políticas, legais, sociais e administrativas entre o Estado e os cidadãos, cujos elementos semânticos são submetidos a fórmulas pré-estabelecidas” (BELLOTTO, 2011). A finalidade da Diplomática em seu nascimento é “fornecer a descrição e a explicação da estrutura formal dos atos escritos, sua autenticidade e sua fidedignidade” (BELLOTTO, 2008, p.1).

Entre o final do século XIX e início do século XX, registra-se a chamada “crise diplomática”, que ocorre devido ao empobrecimento das fontes documentais e a falta de segurança quanto ao objeto de estudos da disciplina. As discussões em torno da ampliação deste objeto se estenderam por quase um século. Frente a esta crise, um grupo de intelectuais da École des Chartes, representados, principalmente por Bautier e Tessier, entendem uma evolução do objeto da Diplomática para o documento como produto de uma sociedade concreta.

Ainda no século XIX, a Arquivística surge como disciplina científica no cenário da Revolução Francesa. Com a divulgação do princípio da proveniência em 1841, a Arquivística se dá como disciplina específica para o tratamento dos documentos de arquivo. Mas somente com a publicação do Manual dos Arquivistas Holandeses em 1898, é que a mesma se eleva ao *status* de disciplina científica e independente.

Já no século XX, a partir dos estudos de Hilary Jenkinson e Giorgio Concetti, consagrados teóricos da Arquivística, que o desenvolvimento do objeto da Diplomática encontra sua mais perfeita aproximação ao relacionar-se com a Arquivística no âmbito do documento de arquivo.

[...] é aquele que, quaisquer que sejam sua data, sua forma e seu suporte, foi naturalmente criado ou recebido por uma pessoa física ou jurídica, pública ou privada, num dado momento, no decorrer e para o exercício de suas atividades habituais (DELMAS, 2010, p.129).

No encontro entre a Diplomática e a Arquivística considera-se que seu objeto se amplia em relação ao tempo e espaço seguindo na direção da gênese e contextualizando a produção dos documentos modernos. Bellotto, no Brasil, é uma das autoras que mais discute esta temática. Para ela, a chamada Diplomática Contemporânea, termo utilizado para identificar este novo momento da disciplina, afasta-se cada vez mais de sua perspectiva

clássica e sua metodologia atual é focada “no estudo dos tipos documentais e de toda sua vinculação orgânica de gênese, atuação e uso” (BELLOTTO, 2011, s/p).

A análise das competências, funções e atividades que se articulam nos procedimentos administrativos foi incorporada ao método da Diplomática Contemporânea para analisar o documento de arquivo, ou seja, o tipo documental. E sua finalidade é desvendar o vínculo arquivístico, isto é, a proveniência e a organicidade do documento.

O uso desta metodologia aplicada aos arquivos tem mostrado muitas vantagens no desenrolar das atividades e rotinas arquivísticas, como afirma Bellotto (2008, p.76), principalmente no que tange à classificação/arranjo por facilitar o entendimento da composição das séries.

Diante do exposto, é nos estudos da identificação, momento arquivístico onde se realizam os estudos de órgão produtor e dos tipos documentais, utilizando-se da metodologia da Diplomática e da Tipologia Documental, que podemos verificar sua aplicabilidade em programas de gestão de documentos.

4 IDENTIFICAÇÃO ARQUIVÍSTICA

A partir da segunda metade do século XX, principalmente após a Segunda Guerra Mundial, a chamada explosão documental, resultou em problemas que afetaram as administrações públicas pelo mundo inteiro. Em contrapartida, a maneira pela qual se dava o tratamento dessa massa documental acumulada não era suficiente para resolver tal problemática, nem tão pouco, a produção documental era diminuída ao ponto de minimizar o crescimento desordenado dos documentos nos arquivos.

A busca pela cientificidade na área da Arquivística traz à tona a necessidade de fixar seu objeto e seus métodos. A produção desordenada de documentos e o acúmulo das massas documentais nos arquivos tornará eminente a consolidação da gestão documental e o uso de metodologias que resolvessem o problema dos arquivos, contexto no qual surgem os debates sobre a identificação.

Neste ambiente surge o termo identificação, utilizado para designar as tarefas de pesquisas realizadas para resolver o problema posto. Grupos de trabalho se formaram em países ibero-americanos, a fim de propor soluções para os arquivos, de forma coordenada, marcando o desenvolvimento da história dos arquivos e o início de uma tradição arquivística (RODRIGUES, 2008, p.48).

Brasil, Portugal, Espanha e México, ainda na década de 1980, na tentativa de resolver o problema posto, formam grupos de trabalho para discutirem uma realidade comum entre estes países de larga tradição administrativa e arquivística e que as práticas cotidianas da administração não correspondiam a transferências sistemáticas de documentos aos arquivos provocando uma acumulação indiscriminada (CARMONA MENDO, 2004, p.41). De acordo com a publicação dos trabalhos apresentados durante a realização das Primeras Jornadas de la Identificación y Valoración de Fondos Documentales de las Administraciones Publicas, 1991, constata-se que na Espanha, a metodologia da identificação se consolidou definitivamente e é definida como:

processo de investigação e sistematização de categorias administrativas e arquivísticas nas quais se sustenta a estrutura de um fundo, sendo um de seus objetivos principais assegurar através de seus resultados a avaliação das series documentais (CARMONA MENDO, 2004, p.41).

Refletindo sobre a metodologia desenvolvida pelos grupos de trabalho formados na década de 1980, percebe-se o estreito vínculo com o princípio da proveniência o qual é básico para o desenvolvimento das tarefas arquivísticas. A base metodológica que sustenta a identificação é a aplicação de maneira direta do principio de proveniência e da ordem original.

O estudo prévio do tipo documental era imprescindível para a identificação das séries documentais, a fim de implantar programas de gestão documental. Da necessidade de caracterização dos tipos documentais, como salienta Rodrigues (2008, p.55), surge “o debate das relações estabelecidas entre a arquivística e a diplomática”.

QUADRO 3 – Características da metodologia da identificação

CONCEITO	pesquisa sobre os elementos que caracterizam o contexto e a tipologia documental produzida
OBJETO	o órgão produtor e seus documentos
BASE METODOLÓGICA	Princípio da proveniência e ordem original

Fonte: RODRIGUES, 2008.

A base metodológica que sustenta a identificação é a aplicação de maneira direta do principio de proveniência, que consiste em não misturar os documentos gerados por diferentes órgãos produtores e, conseqüentemente, o da ordem original, que significa não romper a ordem natural de produção dos tipos documentais oriundos dos procedimentos administrativos que o determinaram. A identificação é a primeira fase da metodologia arquivística, do tipo intelectual, que consiste na pesquisa sobre gêneses do fundo, ou seja, “o sujeito produtor e seu objeto produzido” (MARTÍN-PALOMINO; TORRE MERINO, 2000, p.42).

A identificação, uma metodologia analítica, é realizada em duas etapas:

1. *Identificação do órgão produtor* – contexto onde os documentos de arquivo são produzidos. Para tal é necessário, verificar os elementos que o caracterizam.

***Elemento orgânico* – identificar o órgão produtor da documentação, diferenciando-o do remetente quando assim procede, pois pode não coincidir.**

***Elemento funcional* – representado pelas funções e atividades administrativas desempenhadas pelo órgão, em virtude da competência que tem a seu cargo e que compõe a série documental. (MOLINA NORTES; LEYVA PALMA, 1996, p.158).**

2. *Identificação do tipo documental*

Esta fase da pesquisa pressupõe o levantamento de informações sobre os elementos que caracterizam os documentos de arquivo, o tipo documental. O vínculo estreito existente entre o documento e o procedimento que lhe dá origem se revela através do método de análise de suas características (RODRIGUES, 2008, p.73).

Neste momento é necessário o estabelecimento de conceitos que nortearão todo desenvolvimento da parte aplicada da pesquisa: o conceito de série e de tipo documental. A definição correta da série documental torna seguro o tratamento técnico do documento durante todo seu ciclo vital. O Dicionário de Terminologia Arquivística (2010, p.76), define série como “sequência de unidades de um mesmo tipo documental. Este conceito é representado por Rodrigues (2008, p.74), através da seguinte fórmula: “série = sujeito produtor + função + tipo documental”. O mesmo dicionário define tipo documental como: “configuração que assume uma espécie documental, de acordo com a atividade que a gerou” e pode ser representado pela fórmula: tipo documental = espécie + atividade (verbo + objeto da ação).

Para concluir, pode-se agora estabelecer as relações encontradas entre a identificação arquivística e as outras funções da área. Assim, a identificação fornece as bases para o desenvolvimento de todo o tratamento arquivístico nas várias etapas do ciclo de vida dos documentos, principalmente para a classificação.

5 CLASSIFICAÇÃO NA ARQUIVÍSTICA

É importante iniciar essa discussão tentando conceituar o termo classificação. Para Piedade (1983, p.16), “classificar é dividir em grupos ou classes, segundo as diferenças e semelhanças. Significa dispor os conceitos, segundo suas semelhanças e diferenças, em certo número de grupos metodicamente distribuídos.” Essa definição traz nela, as principais

características do ato de classificar, ou seja, uma ação organizadora que agrupa um conjunto de seres ou objetos por suas semelhanças.

A classificação é considerada como uma das operações técnicas que objetiva dar acesso, estabelecer o contexto da produção documental e fundamentar a avaliação e a descrição no âmbito da gestão de documentos. Seu conceito não é estranho à Arquivística, sua definição segundo o Dicionário de Terminologia Arquivística, 2010, é a “sequência de operações que, de acordo com as diferentes estruturas, funções e atividades da entidade produtora, visam a distribuir os documentos de um arquivo”. Porém, o que soa como estranho nesta atividade é a falta de um fundamento teórico-metodológico nos instrumentos de classificação da área. Para Indolfo (2008, p.55),

“só a classificação permite a compreensão do conteúdo dos documentos de arquivo dentro do processo integral de produção, uso e acesso à informação arquivística, mantendo os vínculos orgânicos específicos que possui com a entidade geradora.”

Para falar de Classificação na Arquivística, é necessário buscar as bases fundamentais das atividades arquivísticas e que está ancorada no princípio da proveniência e da ordem original. O princípio da proveniência, de acordo com Herrera (2003, p.3-4), indica que:

Assim, quando nos enfrentamos com o princípio da proveniência temos que reconhecê-lo como primeiro, principal, natural e geral princípio da Ciência Arquivística. Configurar toda nossa metodologia, estando presente em todas as intervenções arquivísticas. Dará especificidade aos arquivos distinguindo-os das Bibliotecas e Centros de Documentação. (HEREDIA HERRERA, 2003, p.3-4, tradução nossa).

Analisando o conceito de classificação e os princípios básicos da Arquivística, observa-se que os primeiros caminhos na construção de um instrumento de classificação são obtidos. Para Campos (2001, p.48-49), um “dos primeiros passos na elaboração de uma estrutura classificatória é a definição das unidades que constituem o sistema. Essas unidades, na verdade, representam os conceitos e suas relações”, que para a classificação dos documentos arquivísticos, constituem-se no fundo de arquivo e sua ordem original.

Segundo Schellenberg (2006, p.88), a classificação pode ser dividida em três tipos: funcional, organizacional e por assunto. O processo de elaboração dos instrumentos de classificação não é tarefa das mais fáceis na Arquivística, sua elaboração não compreende apenas a escolha do método a seguir (estrutural, funcional ou por assunto), mas deve compreender uma série de escolhas que o tornarão utilizáveis ou não. “A elaboração de plano de classificação não pode estar desconectada da preocupação com a sua aplicação” (GONÇALVES, 1998, p.).

Para tanto, Renato Tarciso Barbosa de Sousa (2009, p.115-116), estudioso do tema em questão, aponta os requisitos e princípios para a construção de planos de classificação para documentos de arquivo, a saber:

- a classificação é uma representação da realidade e, como tal, uma aproximação. É necessário, entretanto, definir níveis de tolerância quanto à indeterminação. Há limites para essa operação. Limites reconhecidos na própria característica humana de sua elaboração. As indeterminações são mais comprometedoras quando ocorrem nos níveis mais altos da estrutura de classificação;
- a classificação persegue um fim, um objetivo, uma finalidade. No caso dos arquivos, é a organização dos documentos;
- a classificação em arquivos presume um agente classificador, que é o arquivista;
- a classificação tem por trás um mecanismo classificador que executa as operações necessárias, tanto em termos do estabelecimento de uma denominação adequada a cada unidade de classificação quanto nas exigências de proliferação, de derivação, de linearidade e simetria, de circularidade, de hierarquia e subordinação;
- a classificação necessita de um princípio de classificação ou de divisão estabelecido pela finalidade e pelas características e relação dos objetos;
- o princípio de classificação é natural quando é o mais adaptado (aproximado) ao conjunto documental;
- o fundamento da divisão deve se manter inalterado para todas as divisões realizadas no mesmo nível, ou seja, deve-se utilizar um único princípio de classificação de cada vez;
- a relação entre os vários níveis (relação entre os conceitos) não pode se resumir ao modelo aristotélico de gênero/espécie, outras relações podem ser utilizadas na articulação entre os níveis;
- as divisões, resultado da aplicação de um princípio de classificação, não devem ser vazias e nem sobrepor-se no seu conjunto. Devem sempre exaurir a totalidade do domínio (extensão) de cada nível de classificação;
- a classificação em arquivos exerce-se sobre um conjunto finito de elementos: o arquivo de uma pessoa física ou jurídica. Ela provém de um único ponto de partida (conjunto inicial);
- a classificação deve ser lida simultaneamente de cima para baixo e de baixo para cima;

- a classificação em arquivos tem como resultado um produto: o instrumento de classificação (plano de classificação), e é nele que deverão estar representadas as equivalências e as hierarquias.

A falta de um modelo metodológico consistente para a construção de instrumentos de classificação é uma realidade brasileira e a busca por meios que transformem essa realidade tem sido um dos esforços de pesquisa na área Arquivística. O diálogo entre a Arquivística e outras áreas do conhecimento, como a Administração, tem se mostrado possível e necessário para encontrar soluções teóricas e aplicadas.

Para a elaboração de planos de classificação, conceitos como competência, missão, função, atividade, tarefa, atividade-fim e atividade-meio, podem nos levar ao entendimento do sujeito produtor de fundos de arquivo. Sousa (2004, p.193-194), propõe uma pesquisa em Arquivística a fim de conhecer o sujeito produtor de arquivos. Para o autor esta pesquisa se fundamenta em três elementos, a saber:

- a organização em toda sua dimensão histórica: criação, extinção, relacionamento com outras organizações e vinculações hierárquicas;
- a organização na sua individualidade: a missão, a estrutura, as funções, atividades e os procedimentos formais e informais;
- os documentos acumulados (produzidos e/ou recebidos): contextualização quanto às condições de sua produção e acumulação. As tipologias documentais como decorrência natural das funções atribuídas a uma organização ou entidade.

É do somatório entre as espécies produzidas pelo exercício das atividades de uma organização, que temos como resultado o tipo documental, objeto a ser classificado através de um plano de classificação.

6 IDENTIFICAÇÃO ARQUIVÍSTICA APLICADA À ELABORAÇÃO DE PLANO DE CLASSIFICAÇÃO: UM ESTUDO PARA A ÁREA DE PESSOAL DA UFRJ

O objeto de coleta e análise de dados para a realização do estudo de caso é a Pró-Reitoria de Pessoal (PR4) que corresponde atualmente à área administrativa da UFRJ, responsável pela gestão de pessoal da universidade, ou seja, abrange as atribuições de gestão e de desenvolvimento do servidor. O levantamento de dados iniciou-se pela busca da legislação pertinente de criação e funcionamento, tanto da universidade, quanto da pró-reitoria, resultando no repertório legislativo embasado no modelo dos autores Molina Nortes e Leyva Palma (1996, p.159), considerando apenas dois elementos: a legislação e o ano, como no quadro abaixo:

QUADRO 4 – Repertório Legislativo

Legislação	Ano
Decreto nº 14.343 – cria a URJ	1920
Lei nº 452 – organiza a UB	1937
Lei nº 4831– cria a UFRJ	1965
Decreto nº 66.536 – aprova o Estatuto da UFRJ	1970
Regimento da Reitoria da UFRJ	1970
Normas de funcionamento da DVRH/PR4	2012
Normas de funcionamento da DVLE/PR4	2011
Normas de funcionamento da DVRB/PR4	2011/2012
Lei nº 8.112 – RJU	1990

Fonte: Elaboração da autora.

Historicamente, a legislação de criação da UFRJ sofreu algumas transformações. Cabe lembrar que sua institucionalização se deu na década de 1920, como Universidade do Rio de Janeiro (URJ). Na década de 1930 passa para Universidade do Brasil (UB) e mais tarde, na década de 1960, se transforma em Universidade Federal do Rio de Janeiro. Nos anos de 1970, a UFRJ publica seu Estatuto de Regimento, que apesar de ultrapassados continuam vigentes. As normas de funcionamento citadas acima ainda não estão vigentes, pois se encontram em fase de revisão para publicação, mas, constituíram uma grande fonte de informação, cedidas pelos gestores das áreas administrativas no momento das entrevistas.

O estatuto da universidade prevê que cada unidade acadêmica ou administrativa elabore seu regimento próprio, no caso da PR4 está inserida na Administração Geral, seu regimento é o da Reitoria vigente desde 1970. Neste regimento, o que hoje é uma pró-reitoria, antes era apenas uma superintendência, denominada Superintendência Geral de Pessoal e Serviços Gerais (SG4), agregando não só as atividades de pessoal, mas também as atividades de serviços gerais e de assistência e segurança, que atualmente compõe a Pró-Reitoria de Gestão e Governança (PR6). Devido à obsolescência do regimento foi necessário elaborar uma equivalência entre as antigas áreas administrativas com as atuais para que o estudo de órgão produtor pudesse ser realizado.

QUADRO 5 - Equivalência de áreas administrativas entre os organogramas de 1970 e atual

ORGANOGRAMA 1970	ORGANOGRAMA ATUAL
SUPERINTENDÊNCIA GERAL DE PESSOAL E SERVIÇOS GERAIS – SG4	PRÓ-REITORIA DE PESSOAL - PR4
-----	SUPERINTENDÊNCIA
-----	<i>1 Divisão de Pessoal</i>
Divisão de Pessoal (com supressão da função Treinamento)	Coordenação de Gestão de Pessoal
<i>Seção de Assentamentos e Cadastro</i>	Subcoordenação de Cadastro
<i>Seção de Direitos e Deveres</i>	<i>2 Divisão de Legislação (DVLE)</i>
	<i>2.1 Seção de Direitos e Deveres (SEDD)</i>

	2.2 <i>Seção de Aposentados</i> (SEIN)
	2.3 <i>Seção de Análise e Procedimentos Disciplinares</i> (SEPD)
	2.4 <i>Seção de Pensionistas</i> (SEPS)
<i>Seção de Classificação de Cargos</i> <i>Seção de Movimentação</i> <i>Seção de Seleção e Treinamento</i>	3 <i>Divisão de Recursos Humanos</i> (DVRH)
	3.1 <i>Seção de Planejamento e Acompanhamento Admissional</i> (SEPA)
	3.2 <i>Seção de Movimentação de Pessoal</i> (SEMP)
	3.3 <i>Seção de Seleção</i> (SELE)
<i>Seção de Controle</i>	4 <i>Divisão de Remuneração e Benefícios</i> (DVRB)
	4.1 <i>Seção de Análise de Pagamento</i> (SEAP)
	4.2 <i>Seção de Benefícios</i> (SEBE)
	4.3 <i>Seção de Pagamento de Pessoal</i> (SEPP)

Fonte: Elaboração da autora.

Observa-se neste momento, que o crescimento da área é visível. O que estava no nível de superintendência foi abarcado com a criação da pró-reitoria, mas a superintendência não deixa de existir, ela assume um nível abaixo. O que estava no nível de divisão passou ao nível de coordenação e o que estava no nível de seção passou ao nível de divisão, acarretando na divisão de suas atribuições. Além disso, percebe-se a criação de outros níveis administrativos como coordenação e subcoordenação não previstas no regimento, bem como a criação de novas seções subordinadas às divisões. Cabe ainda salientar, que a atual Divisão de Recursos Humanos abarcou as atividades de três antigas seções: a Seção de Classificação de Cargos, Seção de Movimentação e Seção de Seleção e Treinamento com supressão da atividade treinamento. Esta atividade prevista na Divisão de Pessoal e realizada pela Seção de Seleção e Treinamento do regimento, hoje faz parte da Divisão de Desenvolvimento de Pessoal (DVDE) que está inserida na Coordenação de Desenvolvimento de Pessoal e que não se constitui como objeto deste estudo.

Com a análise das equivalências de áreas administrativas, pode-se neste momento, elaborar uma amostragem de áreas que foi o objeto de estudo de órgão produtor e que serviu de base para a construção do plano de classificação a saber: a Pró-Reitoria como um todo, a Superintendência Geral de Pessoal, a Coordenação de Gestão de Pessoal, a Divisão de Recursos Humanos com as Seções de Planejamento e Acompanhamento Admissional e a Seção de Movimentação de Pessoal, a Divisão de Legislação com a Seção de direitos e Deveres, a Seção de Aposentados, a Seção e Análise e Procedimentos Disciplinares e a Seção de Pensionistas, e a Divisão de Remuneração e Benefícios com a Seção de Análise de Pagamento, a Seção de Benefícios e a Seção de Pagamento de Pessoal.

A fim de exemplificar o estudo das atribuições do órgão produtor, segue abaixo quadro com uma amostragem do levantamento realizado em uma das áreas da Pró-Reitoria de Pessoal.

QUADRO 6 - Identificação de Órgão Produtor: estudo das atribuições da PR4 (2013)

PRÓ-REITORIA DE PESSOAL – PR4				
MISSÃO: Gestão de Pessoal da UFRJ				
SUPERINTENDÊNCIA GERAL DE PESSOAL				
COMPETÊNCIA: Gerir de forma concisa as atribuições da Coordenação de Gestão de Pessoal				
COORDENAÇÃO DE GESTÃO DE PESSOAL				
FUNÇÃO1- Coordenar as atividades de classificação e redistribuição de cargos; 2- coordenar as atividades de recrutamento e seleção por concurso da universidade; 3- coordenar a movimentação do servidor desde a sua entrada na universidade.				
Área Administrativa	Atividade	Área Administrativa	Tarefa	Tipo Documental
Divisão de Recursos Humanos - DVRH	Promover e planejar a admissão de pessoal	Seção de Planejamento e Acompanhamento Admissional – SEPA	Nomear e empossar os docentes no quadro único da UFRJ Nomear e empossar os técnico-administrativos no quadro único da UFRJ	Processo de nomeação de docentes Processo de nomeação de técnico-administrativos

Fonte: Elaboração da autora.

A partir da conclusão do estudo das atribuições do órgão produtor, neste caso, a PR4, é necessária para identificar e nomear corretamente o objeto a ser classificado, ou seja, o tipo documental. Abaixo, segue quadro com amostragem dos tipos documentais encontrados na Seção de Planejamento e Acompanhamento Admissional (SEPA).

QUADRO 7 – Denominação do tipo documental

Área Administrativa (2013)	Tipo Documental (2013)	Denominação do Tipo Documental (espécie+verbo+objeto da ação)
SEPA	Processo de nomeação de docente	Processo de nomeação de docente

Fonte: Elaboração da autora.

Encerrando o estudo de órgão produtor, passa-se neste momento, para elaboração da proposta de plano de classificação que foi desenvolvido a partir de uma amostragem de áreas administrativas da PR4 da UFRJ, com o objetivo de validar a metodologia da identificação

como primeira etapa das atividades arquivísticas, embasando a classificação e a avaliação. O estudo de órgão produtor leva ao conhecimento geral das áreas administrativas e suas atribuições. Diante deste conhecimento podemos representá-las através do plano de classificação, que é uma aproximação que reflete o contexto de produção dos documentos.

Nesta perspectiva, consideraram-se os requisitos e princípios propostos por Sousa (2009, p.115-116), para a construção de planos de classificação, conforme citado anteriormente.

Na UFRJ existe um sistema eletrônico para acompanhar o andamento dos processos abertos pela universidade, chamado Sistema de Acompanhamento de Processo (SAP). Neste sistema todas as unidades da UFRJ estão representadas por um código numérico, que reflete a hierarquia entre os órgãos, o que é facilmente percebida por seus usuários. Por ser uma codificação conhecida na universidade, foi utilizada para representar a subordinação das classes no plano de classificação. Segue um quadro com os códigos das áreas da PR4 usadas neste estudo:

QUADRO 8: Código das áreas administrativas da PR4

Área Administrativa	Código	Atribuição
Pró-Reitoria de Pessoal	6400	Missão
Superintendência Geral de Pessoal	6403	Competência
Coordenação de Gestão de Pessoal	6410	Função
DVLE	6411	Atividade
SEDD	641110	Tarefa
SEIN	641120	Tarefa
SEPD	641130	Tarefa
SEPS	641140	Tarefa
DVRH	6412	Atividade
SEPA	641220	Tarefa
SEMP	641230	Tarefa
SELE	641240	Tarefa
DVRB	6413	Atividade
SEAP	641310	Tarefa
SEBE	641320	Tarefa
SEPP	641330	Tarefa

Fonte: Sistema de Acompanhamento de Processo (SAP)

A partir da utilização destes códigos, têm-se as seguintes explicações:

6410.1

Identificação da área de coordenação
Numeração correspondente à função

6410.1.6411.1

Identificação da área de coordenação
 Numeração correspondente à função
 Identificação da área de divisão subordinada à coordenação
 Numeração correspondente à atividade

641110.1

Identificação da área de seção subordinada à divisão
 Numeração correspondente à tarefa

Após as explicações pertinentes, segue a apresentação de uma amostra do modelo de plano de classificação funcional realizado a partir da identificação de órgão produtor e suas atribuições.

QUADRO 9 - Plano de Classificação Funcional para a área de pessoal da UFRJ (uma amostragem)

6400 MISSÃO: Gestão de Pessoal da UFRJ

6403 COMPETÊNCIA: Gestão das atribuições da Coordenação de Gestão de Pessoal

6410.2 Coordenação das atividades de recrutamento e seleção de pessoal

6410.2.6412.1 Promoção e planejamento da admissão de pessoal

641220.1 Nomeação e posse dos docentes

Tipo documental: Processo de nomeação de docente

641220.2 Nomeação e posse dos técnico-administrativos

Tipo documental: Processo de nomeação de técnico-administrativos

Fonte: Elaboração da autora.

Os procedimentos realizados para a elaboração deste plano de classificação foram:

1. Levantamento das fontes legislativas (leis, decretos, estatutos, regimentos, organogramas, etc.);
2. Estudo da história administrativa da universidade a fim de contextualizar o órgão produtor: a PR4 da UFRJ;
3. Atualização dos dados: se as fontes levantadas não forem suficientes para a composição do estudo de elemento funcional, é necessária a realização de entrevistas com gestores (caso de fontes muito antigas);
4. Elaboração de uma equivalência de áreas administrativas a fim de compatibilizar dados antigos com atuais. (Se o organismo tiver suas fontes legislativas atualizadas, ignorar esta etapa);

5. Estudo do elemento funcional a fim de verificar as atribuições do órgão estudado;
6. Estudo dos tipos documentais no primeiro nível a fim de denominar as séries documentais;
7. Identificação dos níveis necessários para dividir internamente as classes e subclasses;
8. Denominação das classes e subclasses obedecendo as redações das atribuições das áreas administrativas contidas nas fontes de informação (competência, função, atividade, tarefa), representando sua hierarquia. Recomenda-se cortar as explicações e os termos excessivos que constam na redação das fontes informacionais, utilizando-se da fórmula: verbo(s) + matéria da ação;
9. Codificação das classes e subclasses. Se o organismo estudado tiver alguma codificação que represente as áreas administrativas, recomenda-se observar se esta pode ser também utilizada para a representação das classes e subclasses;
10. Validação do plano de classificação com o gestor da área.

7 CONSIDERAÇÕES FINAIS

A análise dos princípios teóricos e metodológicos que fundamentam a construção de planos de classificação na perspectiva da Arquivística brasileira, bem como os resultados obtidos no âmbito dos arquivos universitários, especificamente na área de pessoal da UFRJ constituiu-se no foco desta pesquisa.

No conceito de identificação está imbuída a necessidade de investigação do sujeito criador/acumulador de arquivos e seus produtos, ou seja, os documentos. Este processo se concretiza com estudo de identificação de órgão produtor e de tipos documentais, ancorados no princípio da proveniência e na ordem original, sendo precípuo da classificação e da avaliação dos documentos.

O processo de elaboração de instrumentos de classificação não é tarefa das mais fáceis na Arquivística. Sua elaboração não compreende apenas a escolha do método a seguir (estrutural, funcional ou por “assunto”), mas deve compreender uma série de escolhas que o tornarão utilizáveis ou não. Neste sentido é preciso considerar a forma como os usuários usam e buscam os documentos a fim de complementar o instrumento de classificação.

Buscou-se com esta pesquisa refletir sobre os aspectos de construção de planos de classificação arquivísticos sob uma ótica científica, analisando critérios sólidos embasados na investigação, tentando contribuir para o desenvolvimento da Arquivística brasileira como Ciência, *status* que lhe é merecedor. De fato, a pesquisa aqui apresentada não encerra, de

modo algum, as discussões sobre o tema, pelo contrário, incita cada vez mais a busca por novos referenciais que irão elevar o conhecimento arquivístico.

REFERÊNCIAS

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS (ABNT). Disponível em: <www.abnt.org.br>. Acesso em 29 abr. 2013.

BELLOTTO, Heloísa Liberalli. **Diplomática e tipologia documental em arquivos**. 2. ed. Brasília, DF: Briquet de Lemos, 2008. 106 p.

BELLOTTO, Heloísa Liberalli. Diplomática para quê? In: CONGRESSO BRASILEIRO DE PALEOGRAFIA E DIPLOMÁTICA, 1., 2011, Campos dos Goytacazes. **Apresentação...** Rio de Janeiro, 2011.

BRASIL. Constituição (1988). **Constituição Federal do Brasil**. São Paulo: Editora Escala, 2007.

BRASIL. **Estatuto da Universidade Federal do Rio de Janeiro**. Disponível em:<http://www.ufrj.br/pr/conteudo_pr.php?sigla=ESTATUTO>. Acesso em 10 abr. 2012.

CAMPOS, Maria Luiza de Almeida. **Linguagem documentária: teorias que fundamentam sua elaboração**. Niterói: EdUFF, 2001. 133 p.

CARMONA MENDO, Concepción. Consideraciones sobre el método em archivística. **Documenta & Instrumenta**, Madrid, v. 1, p. 35-46, 2004. Disponível em:<<http://www.ucm.es/info/documen/htm/default.htm>>. Acesso em: 21 ago.2011.

DELMAS, Bruno. **Arquivos para quê?** textos escolhidos. Tradução de Danielle Ardaillon. São Paulo: Instituto Fernando Henrique Cardoso, 2010.

DICIONÁRIO brasileiro de terminologia arquivística. Rio de Janeiro: Arquivo Nacional, 2005.

DICIONÁRIO de terminologia arquivística. São Paulo: Centro de Memória da Educação, 2010.

GONÇALVES, Janice. **Como classificar e ordenar documentos de arquivo**. São Paulo: AAB-SP / Arquivo do Estado, 1998. (Projeto como fazer).

GRUPO DE TRABAJO DE LOS ARCHIVEROS MUNICIPALES DE MADRID. Identificación y valoración de fondos documentales municipales: la experiencia del Grupo de Trabajo de Archiveros Municipales de Madrid. In: PRIMERAS JORNADAS SOBRE METODOLOGÍA PARA LA IDENTIFICACIÓN Y VALORACIÓN DE FONDOS DOCUMENTALES DE LAS ADMINISTRACIONES PUBLICAS, 1991. **Actas...** Madrid: Dirección de Archivos Estatales,1992. p. 491-496.

HERRERA, Antonia Heredia. **El principio de procedência y los otros principios de La Archivística**. São Paulo: Associação dos Arquivistas de São Paulo, 2003. 14 p.

INDOLFO, Ana Celeste. **O uso das normas arquivísticas no Estado brasileiro: uma análise do Poder Executivo Federal**. 2008. Dissertação (Mestrado)– Convênio Universidade Federal Fluminense e Instituto Brasileiro de Informação em Ciência e Tecnologia, 2008. Disponível em:<<http://ibict.phlnet.com.br/anexos/anacelesteindolfo.pdf>> Acesso em: 12 set. 2012.

MARCONDES, Carlos Henrique. Arquivologia, Biblioteconomia e Documentação: o lugar da Ciência da Informação entre as Ciências da Documentação. **Revista EDICIC**, v. 1, n. 4, out./dez. 2011. Disponível em: <<http://www.edicic.org./revista/>>. Acesso em: 23 jan. 2012.

MARTÍN-PALOMINO BENITO, M.; TORRE MERINO, J. L. L. **Metodología para la identificación y valoración de fondos documentales**. Madrid: Ministerio de Educación, Cultura e Deporte, 2000. 109 p. (Escuela Iberoamericana de Archivos: Experiencias y materiales).

MOLINA NORTES, Juana; LEYVA PALMA, Victoria. **Técnicas de archivo y tratamiento de la documentación administrativa**. Guadalajara: ANABAD Castilla – La Mancha, 1996. (Textos de ANABAD Castilla – La Mancha, 2).

PIEIDADE, M. A. Requião. **Introdução à Teoria da Classificação**. 2. ed. Rio de Janeiro: Interciência, 1983.

RIO DE JANEIRO (Estado). Governo do Estado do Rio de Janeiro. **Manual de Gestão de Documentos**. Rio de Janeiro, 2012. 108 p.

RODRIGUES, Ana Célia. **Diplomática contemporânea como fundamento metodológico para a identificação de documentos de arquivo**. 2008. Tese (Doutorado)– Universidade de São Paulo, São Paulo, 2008. Disponível em: <www.teses.usp.br/teses/.../tde-27112008-151058/>. Acesso em: 17 set. 2011.

SCHELLENBERG, T. R. **Arquivos modernos: princípios e técnicas**. Tradução de Nilza Teixeira Soares. 6. ed. Rio de Janeiro: Editora FGV, 2006. 388 p.

SOUSA, Renato Tarciso Barbosa de. A classificação como função matricial do que- fazer arquivístico. In: SANTOS, Vanderlei Batista dos; INNARELLI, Humberto Celeste; SOUSA, Renato Tarciso Barbosa de (Org.). **Arquivística: temas contemporâneos: classificação, preservação digital, gestão do conhecimento**. 3. ed. Brasília, DF: SENAC, 2009. p. 79-172.

SOUSA, Renato Tarciso Barbosa de. **Classificação em arquivística: trajetória e apropriação de um conceito**. 2004. Tese (Doutorado)– Universidade de São Paulo. São Paulo, 2004. Disponível em: <www.teses.usp.br/teses/.../>. Acesso em: 17 set. 2011.

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO. Divisão de Gestão Documental e da Informação. Disponível em: <www.dgdi.ufrj.br>. Acesso em: 8 abr. 2013.

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO. Regimento Interno da Reitoria. **BUFRJ**, Rio de Janeiro, n. 51, 1970.

ANÁLISE E DESCRIÇÃO DOS METADADOS PRESENTES NOS ELEMENTOS DA MODELIZAÇÃO DE UM DOMÍNIO ORGANIZACIONAL

DESCRIPTION AND ANALYSIS OF METADATA ELEMENTS PRESENT IN THE MODELING OF AN ORGANIZATIONAL DOMAIN

Célia da Consolação Dias
Lídia Alvarenga

Resumo: Os objetivos deste trabalho são descrever os metadados identificados no sistema categorial organizado a partir da análise dos formulários usados na Metodologia DIRKS para a elaboração dos instrumentos de gestão arquivística na administração direta do Estado de Minas Gerais. A modelagem partiu das categorias fundamentais de Ranganathan. Foram analisados oito instrumentos de coleta de dados, usados na implantação do projeto de elaboração dos instrumentos de gestão de documentos do governo do Estado de Minas Gerais. A metodologia deste estudo compreendeu doze etapas, desde a separação das entidades de cada formulário até a elaboração de um sistema categorial representando um domínio organizacional, composto de categorias, subcategorias e subdivisões de subcategorias, bem como, a descrição dos metadados presentes nos formulários usados para a elaboração dos instrumentos de gestão de documentos arquivísticos. A metodologia usada nesta pesquisa e as decorrentes reflexões evidenciam os princípios teóricos da área: os princípios do respeito aos fundos, da ordem original e a abordagem da teoria das três idades que fundamentaram a interpretação dos elementos dos formulários de coleta de dados.

Palavras-chave: Análise de domínio e arquivística. Modelagem organizacional. DIRKS - *Designing and Implementing Recordkeeping Systems*. Metadados arquivísticos. Instrumentos de gestão.

Abstract: The objectives of this study are to describe the metadata identified in organized categorical system from the analysis of the forms used in the DIRKS methodology for the preparation of instruments of records management in the direct administration of the State of Minas Gerais. The domain analysis presented has been oriented by Ranganathan's fundamental categories scheme. There has been analyzed eight (08) application forms, used for the Minas Gerais Brazilian State Government documents management system planning and development. As theoretical foundations has been discussed domain representation and knowledge modeling concepts This study methodology included twelve (12) steps, from gathering applications forms entities, until designing an organizational model, using categories, subcategories and categories subdivisions, well as a description of the metadata present in the forms used for the preparation of management tools records. The methodology used in this research and the resulting reflections show the theoretical principles of the field: the principles of respect of funds, from the original order and the approach of the theory of the three ages that supported the interpretation of the elements of the data collection forms.

Key-words: Domain analysis and records managements. Archives metadata. DIRKS - *Designing and Implementing Recordkeeping Systems*. instruments of records management.

1 INTRODUÇÃO

Partindo do pressuposto de que a arquivística por sua própria natureza e método tem na análise do domínio organizacional seu fundamento clássico, este artigo apresenta uma análise e a descrição dos metadados presentes nos elementos da modelização de um domínio

organizacional. Tal análise foi feita após a identificação e o agrupamento dos metadados extraídos de formulários integrantes da Metodologia DIRKS, *Designing and Implementing Recordkeeping Systems*, que nesse trabalho será chamada de Metodologia DIRKS. Esta metodologia foi usada para a criação dos instrumentos de gestão de documentos de arquivos na administração direta do Governo do Estado de Minas Gerais.

A interação com o ambiente organizacional e o impacto causado pelo uso das informações oriundas do ambiente interno e externo têm levado as organizações a se preocuparem com o processo de gerenciamento das informações produzidas e recebidas, com vistas ao seu uso para a tomada de decisões. Entretanto um ponto importante para garantir o uso das informações para a tomada de decisões é a implementação das atividades relacionadas à gestão de documentos. E para que isto ocorra os instrumentos de gestão são fundamentais para garantir o suporte a todas as atividades referentes à tramitação, à avaliação e à destinação de documentos de arquivos. Para a construção desses instrumentos que possibilitam a implementação das atividades próprias da gestão de documentos existem desafios que envolvem tanto as comissões responsáveis pela sua criação quanto os gestores de cada órgão da Administração direta do Governo do Estado. Um grande desafio tanto para a administração pública quanto para as empresas privadas tem sido sensibilizar os gestores da importância das ações da gestão de documentos e que, existem instrumentos próprios e legítimos para a implementação dessas ações nas empresas. Quanto a este último ponto, ressalta-se que os documentos na arquivologia estão relacionados a conceitos como prova ou evidência⁴², conservação, pesquisa e informações, resultantes desse domínio e, conseqüentemente, tais aspectos integram o processo de modelagem dos sistemas de informação.

Para a ciência da informação, a organização de informações sempre fez parte do trabalho de permitir o acesso às informações dos diversos acervos, presentes nas bibliotecas, arquivos e museus e que, estão disponíveis em mídia impressa, eletrônica ou em outro suporte. São essas as informações que as organizações possuem e precisam ter acesso para realizar as suas atividades do dia-a-dia e, em alguns casos, garantir que no futuro elas estarão acessíveis devido ao seu valor de pesquisa. Permitir o acesso a informações significa disponibilizar os dados extraídos dos diversos tipos de documentos e possibilitar ao usuário identificar informações contidas nos metadados relacionados aos seus elementos físicos e de

42 O termo "evidência ou "prova" na forma de registros ajuda a garantir que indivíduos e as organizações são responsáveis perante o governo, tribunais de justiça, acionistas, clientes, comunidade e grupos de interesse das gerações futuras. Austrália (2007, p.4).

conteúdo, para permitir que as pessoas sejam conduzidas até o documento desejado para satisfazer suas necessidades de informação.

Os objetivos deste trabalho são: descrever os metadados identificados no sistema categorial organizado a partir da análise dos formulários usados na Metodologia DIRKS para a elaboração dos instrumentos de gestão arquivística na administração direta do Estado de Minas Gerais.

2 A REPRESENTAÇÃO NA CIÊNCIA DA INFORMAÇÃO

A interação com o ambiente organizacional e o impacto causado pelo uso das informações oriundas do ambiente interno e externo tem levado as organizações a se preocuparem com o processo de gerenciamento das informações geradas e recebidas para agregar valor àquelas já existentes com vistas à sua aplicação e reutilização. Um dos grandes desafios para os profissionais que projetam os sistemas de informação é a modelagem destes sistemas, o entendimento do modelo organizacional, as necessidades dos usuários e o entendimento do próprio domínio do conhecimento que está sendo modelado.

Um modelo organizacional é uma representação explícita e simplificada da organização. O seu principal objetivo é obter uma visão da organização aspecto importante para a Ciência da Informação.

Para a ciência da informação a representação sempre fez parte do trabalho de permitir o acesso às informações das diversas coleções presentes nas bibliotecas, acervo documental dos arquivos e informações sobre os objetos presentes nos museus e que, estão disponíveis em mídia impressa, eletrônica ou outro suporte do contexto digital que as organizações possuem e precisam manusear no dia-a-dia das suas atividades. Permitir o acesso significa disponibilizar os dados extraídos dos diversos tipos de documentos e possibilitar ao usuário identificar os seus elementos físicos e de conteúdo, para que as pessoas sejam conduzidas até o material desejado para satisfazer a sua consulta.

Para Miranda (1999) a representação, no âmbito da Ciência da Informação, visa a promover o acesso ao conteúdo dos documentos para uso e posterior geração de novos conhecimentos.

Outro desafio identificado na literatura tem sido em relação à descrição das informações com vistas a favorecer a recuperação e ao uso posterior do acervo de conhecimento organizado. Como dito anteriormente o desafio tem sido a recuperação da

informação e, como levantado por Mooers⁴³ uma das questões surgidas em torno deste ponto é: “como descrever intelectualmente a informação?” Ou como apontado por Vickery (1986, p.145) em todos os estudos onde a representação do conhecimento é estudada a preocupação tem sido em torno da decisão de “como o conhecimento pode ser representado e que as representações podem ser manipuladas”.

Da mesma forma Café e Bräscher (2008) afirmam que a descrição da informação é fundamental na recuperação da informação, pois cumpre a função de possibilitar o acesso a itens específicos que se encontram armazenados. Os metadados são recursos reconhecidos como fundamentais para a recuperação da informação na web e facilitam gerenciar e recuperar informação, esteja na forma eletrônica ou não. Com isso, de acordo com (Gilliland-Swetland (2002), citado por Lourenço (2005) os metadados têm a função de certificar a autenticidade do conteúdo; estabelecer e documentar o contexto do conteúdo; identificar e explorar as relações estruturais dos objetos de informação externas e internas com outros objetos de informação e fornecer pontos de acesso para os usuários recuperarem estes objetos informacionais.

A preocupação com a descrição do conteúdo tem sido cada vez mais urgente dado o caráter da quantidade de registros agora em meio eletrônico disponível em todo o mundo tal como apontado por Campos & Gomes que, afirmam que “as formas de representação/organização dos domínios de conhecimentos, neste caso, vão implicar diretamente os processos de transferência da informação, que hoje estão diretamente ligados a recuperação em meios eletrônicos.”(CAMPOS & GOMES, 2003)

Ressalta-se que a preocupação com a descrição da informação está relacionada tanto ao entendimento do metadado quanto da necessidade de se trabalhar com padrões que podem contribuir para o entendimento da descrição acerca dos objetos. Nesse caso diz respeito não apenas à descrição dos campos dos formulários, mas também do entendimento de cada um dos elementos no processo de criação dos próprios instrumentos de gestão de documentos possibilitado pela implementação da metodologia DIRKS. Nesse sentido Garshol (2004) afirma que os metadados podem ser entendidos como "qualquer afirmação sobre um recurso de informação",

No contexto de organização da informação e do conhecimento os metadados podem ser usados para descrever, localizar, usar ou gerenciar, para explica um recurso de informação entre outras funções relacionadas à recuperação da informação. Especificamente no domínio

43 Autor citado por Saracevic em seu artigo Ciência da informação:origem, evolução e relações

organizacional e a implementação da metodologia DIRKS observou-se a existência de uma grande quantidade de instrumentos para a realização da coleta de dados. Esses dados compreenderam informações relevantes para a construção de um conjunto de requisitos funcionais de uma organização específica. Para isso foi feito um exaustivo levantamento nas fontes documentais para buscar os elementos sobre o ambiente organizacional no tocante aos aspectos regulatórios, culturais, políticos e administrativos. Neste caso a descrição dos metadados poderia auxiliar o entendimento das comissões de avaliação de documentos no uso dos instrumentos de coleta de dados, durante a implementação de cada passo da metodologia, bem como lidar com a complexidade tanto da metodologia quanto do próprio ambiente organizacional.

3 O DOMÍNIO ORGANIZACIONAL E A METODOLOGIA DIRKS PARA A CONSTRUÇÃO DOS INSTRUMENTOS DE GESTÃO

Um domínio pode ser representado por um campo do conhecimento, uma comunidade de prática ou uma instituição corporativa associada ao um negócio específico, por exemplo. Nesse estudo o domínio analisado compreendeu um domínio organizacional de empresas públicas da administração direta formada por seus processos organizacionais, atores sociais, pela cultura organizacional, pela estrutura da organizacional, funções, atividades, fontes de informação elementos do ambiente, documentos produzidos e recebidos. Para implementar as atividades referentes a essa gestão de documentos tornou-se necessário a criação dos instrumentos de gestão de documentos arquivísticos para a administração pública direta do Governo de Minas Gerais. A Metodologia DIRKS foi usada como suporte para a criação dos dois instrumentos de gestão de documentos arquivísticos.

A Metodologia DIRKS foi desenvolvida pelo Arquivo Nacional da Austrália e está registrada no Manual DIRKS - *Designing and Implementing Recordkeeping Systems*. Esta metodologia tem como objetivo fornecer orientações sobre as melhores práticas para o gerenciamento de informações e documentação, tanto em organizações públicas quanto privadas. Outro ponto importante relacionado a esta metodologia é que, ela tem como base as necessidades do negócio da organização e por isto, possibilita um estudo minucioso das atividades, processos, funções e necessidades de cada órgão, e o conhecimento de tais informações contribuiu para facilitar entendimento do processo de construção dos instrumentos de gestão.

A metodologia DIRKS foi concebida em oito (08) etapas. Esses oito passos compreendem todo o processo de gestão de documentos, da pesquisa sobre a organização,

suas funções e procedimentos, até a avaliação da implementação e a revisão dos instrumentos e do sistema de arquivamento. A metodologia foi organizada da seguinte forma:

- 1) Passo A - Investigação preliminar;
- 2) Passo B - Análise da atividade empresarial;
- 3) Passo C - Identificação dos requisitos de arquivamento;
- 4) Passo D - Avaliação dos sistemas de gestão de documentos arquivísticos existentes;
- 5) Passo E - Identificação de estratégias para gestão de documentos arquivísticos;
- 6) Passo F - Concepção de um sistema para gestão de documentos arquivísticos;
- 7) Passo G - Implementação de um sistema para gestão de documentos arquivísticos;
- 8) Passo H - Avaliação do sistema gestão de documentos arquivísticos após a implementação.

Essa metodologia foi usada no Projeto de Elaboração dos Instrumentos de Gestão das atividades finalísticas, desenvolvido pelo Arquivo Público Mineiro em vinte órgãos da administração direta do governo do Estado de Minas Gerais. O estudo aqui apresentado discute a importância da descrição dos metadados presentes nos formulários da metodologia DIRKS tanto para auxiliar no seu entendimento, quanto para compreender as etapas necessárias para a criação de dois importantes instrumentos de gestão de documentos arquivísticos que, permitirão a implementação das ações de gestão nas organizações. Para isso, entender o esquema de funcionamento dos três primeiros passos da metodologia DIRKS é fundamental.

3.1 Descrição dos Passos A, B e C da metodologia DIRKS

O passo A da metodologia DIRKS, chamado de Investigação Preliminar compreende a elaboração dos levantamentos sobre cada instituição, a estrutura, o ambiente de negócio, de regulação, o quadro sócio-político em que a instituição opera, os fatores que afetam as práticas de arquivamento e outras questões em relação à gestão de documentos e sistemas de arquivamento. O foco dessa etapa é identificar questões sobre arquivamento, problemas específicos ou áreas de riscos, informações sobre a geração de documentos na instituição. O instrumento de coleta de dados busca recolher dados sobre o contexto organizacional, que permite a possibilidade de fazer uma profunda investigação sobre esses elementos e outros que podem estabelecer “as bases para a definição do escopo do projeto de gestão de documentos na instituição.” (Austrália, 2007).

O passo B está relacionado à Análise da Atividade Empresarial e compreende o levantamento de funções e atividades em formulários próprios, bem como à elaboração dos

fluxos de cada processo com a identificação dos documentos. Esses levantamentos compreendem a etapa de pesquisa que corresponde ao levantamento das estruturas administrativas, das funções e das atividades da instituição em questão. A análise de atividades fornecerá um modelo conceitual refletindo o que a organização faz através do levantamento das suas atividades, tal como está preconizado no manual DIRKS.

O objetivo dessa etapa é desenvolver um modelo conceitual do que a organização faz e como o faz através da análise das suas atividades e processos. Esta análise fornecerá base para o núcleo do desenvolvimento de ferramentas gestão de documentos e contribuirá para a tomada de decisões sobre a criação, captura, controle, armazenamento, eliminação e de acesso dos registros nas etapas seguintes. Isto é particularmente importante em ambientes organizacionais digitais em que a conexão entre uma organização e os seus registros é lógica e não física (AUSTRÁLIA, 2007, p.4).

O terceiro Passo C da metodologia DIRKS compreende a Identificação dos Requisitos de Arquivamento. Nessa etapa o objetivo é “identificar e registrar Requisitos da sua organização necessários para manter as evidências das suas atividades e para documentar os requisitos de uma forma estruturada e de fácil manutenção” (AUSTRÁLIA, 2007). Para a operacionalização do passo C é necessário a implementação de duas importantes etapas. Em primeiro lugar fazer a identificação, análise dos requisitos de arquivamento e, em seguida a avaliação dos processos já identificados no passo B e que serão importantes para promover as discussões acerca dos prazos de guarda para a documentação. O entendimento de todos os elementos coletados nos três primeiros passos é fundamental para a construção de um arcabouço representado pelos dados coletados e pela compreensão dos elementos presentes no ambiente corporativo para subsidiar a desenvolvimento dos instrumentos de gestão.

4 PROCEDIMENTOS METODOLÓGICOS

O levantamento das entidades e dos metadados do domínio foi feito nos oito (08) formulários utilizados na implementação dos três primeiros passos da Metodologia DIRKS e que, compreendem o objeto empírico deste estudo. Os seguintes instrumentos de coleta de dados foram analisados: quatro formulários do Passo A, três formulários do Passo B e um do Passo C, conforme detalhado a seguir:

No **Passo A** da Metodologia DIRKS foram usados: Formulário de Identificação de Fontes – Leis e Decretos; Formulário de Identificação de Fontes-Normas Específicas – Regulamentos; Formulário de Identificação de Fontes – Outras Fontes Incluindo Entrevistas e Levantamento do Contexto Organizacional. A etapa chamada, de investigação preliminar permitiu compreender o entendimento da instituição e do contexto administrativo, legal, de

negócios e social no qual ela opera, a avaliação geral dos pontos fortes e fracos da gestão de documentos e dos sistemas de arquivos na instituição.

No **Passo B**: Formulário de Levantamento de Função; chamado de Tabela 1: Análise de Atividades; Análise de atividade - Fluxo da Transação. O objetivo dessa etapa é desenvolver um modelo conceitual do que a organização faz e como o faz através da análise das suas atividades e processos. E, no **Passo C**: o formulário chamado de Tabela 2 - Identificação de Requisitos de Arquivamento.

É importante ressaltar que os elementos analisados nessa pesquisa foram levantados a partir das entidades enunciadas nos títulos das questões ou dos campos presentes em cada formulário, ou seja, dos metadados. Com o objetivo de se obter maiores esclarecimentos foram também consultados exemplares de formulários preenchidos com os dados da Secretaria de Cultura do Estado de Minas Gerais.

4.1 Formação do sistema categorial com os metadados dos formulários

As entidades ou campos propostos para um domínio organizacional, tal como aparecem nos formulários dos Passos A, B e C da Metodologia DIRKS, analisados neste estudo foram organizados em um sistema categorial que representam os metadados da seguinte forma: usando as categorias fundamentais de Ranganathan, com exceção da categoria matéria que não foi aplicável ao domínio estudado, dividindo estas em categorias específicas, subcategorias e divisão das subcategorias. Devido a grande quantidade de entidades resultantes das análises do domínio organizacional, a partir da aplicação da Metodologia DIRKS, neste trabalho serão apresentadas e discutidas apenas os elementos identificados no processo de categorização e que foram usados na construção dos instrumentos de gestão.

As categorias fundamentais extraídas do domínio organizacional, neste estudo, são as quatro categorias de Ranganathan aqui indicadas: 1 Personalidade, 2 Energia, 3 Espaço e 4 Tempo e suas respectivas categorias específicas estão apresentadas, no QUADRO 1.

QUADRO 2 - Categorias Específicas de cada Categoria Fundamental do Sistema Categorial

Da categoria Fundamental PERSONALIDADE 1, temos:	Da categoria Fundamental ENERGIA 2, temos:
<p style="text-align: center;">1.1 Organização</p> <p>1.2 Instrumentos de Apoio</p> <p>1.3 Fonte de informação</p> <p>1.4 Ato Normativo</p> <p>1.5 Projetos, Programa</p>	<p>2.1 Função</p> <p>2.2 Atividade</p> <p>2.3 Transação</p> <p>2.4 Processo</p>

Da categoria Fundamental ESPAÇO 3, temos:	Da categoria Fundamental TEMPO 4, temos:
3.1 Endereço da Unidade	4.1 Ciclo de Vida
3.2 Telefone	4.2 Data do Ato Normativo
3.3 Endereço Eletrônico	4.3 Data do Evento
3.4 Endereço do Site	

Fonte: Elaborado pelas autoras

Para responder às perguntas, propostas no início deste trabalho, e coletar os possíveis elementos para a análise e representação do domínio organizacional, foram realizadas 12 (doze) etapas que compreenderam os procedimentos metodológicos deste estudo, conforme indicado a seguir: Etapa 1: Organização do material empírico; Etapa 2: Análise dos campos dos formulários; Etapa 3: Complementação dos campos dos formulários; Etapa 4: Codificação das entidades dos formulários; Etapa 5: Reunião alfabética das entidades; Etapa 6: Identificação das Categorias Fundamentais; Etapa 7: Agrupamento das entidades para formação das categorias fundamentais; Etapa 8: Identificação das categorias específicas; Etapa 9: Formação das subcategorias; Etapa 10: Extração do conteúdo das entidades; Etapa 11: Montagem do sistema categorial; Etapa 12: Análise e descrição dos metadados de cada categoria.

As entidades propostas para um domínio organizacional, tal como aparecem nos formulários dos três passos da Metodologia DIRKS, foram organizadas em um sistema categorial, apenas para facilitar a análise das mesmas. Neste sentido serão apresentadas a seguir as análises feitas sobre os metadados identificados nos formulários estudados e a sua descrição para auxiliar o processo de construção dos instrumentos de gestão de documentos arquivísticos.

Observou-se que durante o processo de elaboração dos instrumentos de gestão foi importante conhecer o contexto de produção dos documentos, o ambiente regulatório, as atividades realizadas pela instituição e sua relação com os documentos produzidos. Para isto fez-se necessário discutir os elementos e metadados identificados nos formulários da metodologia DIRKS e seu papel no processo de construção dos instrumentos de gestão de documentos arquivísticos.

5 RESULTADOS

Desta forma, apresentam-se a seguir as discussões sobre os metadados identificados durante a análise dos formulários usados para a coleta de dados dos três primeiros passos da metodologia DIRKS. As discussões foram subsidiadas pelo suporte teórico oferecido por autores tanto da arquivologia quanto da ciência da informação. São eles: Thomassen (2006);

CIA (2005); CIA (2008); CIA (2007); ISO15.489; Bernardes (1998); Arquivo Nacional (2005); Metodologia DIRKS apresentada pelo Arquivo Nacional da Austrália; Sheperd & West (2003) e estão organizadas com as análises referentes aos elementos das seguintes categorias fundamentais: 1 Personalidade; 2 Energia; 3 Espaço e 4 Tempo.

Metadados da Categoria fundamental 1 Personalidade

A categoria **1 Personalidade** reuniu as entidades dos formulários que apresentam as características relativas às coisas próprias do domínio organizacional, representada neste estudo por: 1.1 Organização; 1.2 Instrumentos de Apoio; 1.3 Fonte de informação; 1.4 Ato Normativo e 1.5 Projetos, Programas. Observou-se que estas informações foram coletas em diversos formulários adotados nos levantamentos da etapa referente ao passo 1, Investigação preliminar, da Metodologia DIRKS.

Observou-se que a categoria específica **1.1 Organização** da categoria fundamental **1 Personalidade**, contém os elementos que permitem caracterizar a instituição representada. Neste caso foram identificados os elementos da estrutura organizacional fundamentais para a identificação do contexto organizacional onde os documentos foram gerados, bem como a busca de informação para o entendimento do órgão produtor e acumulador dos documentos de cada órgão trabalhado. Essas informações são relevantes, tanto para construir os instrumentos e subsidiar as atividades de gestão de documentos, quanto para identificar processos que serão desenvolvidos nos arquivos permanentes, tais como a descrição arquivística, a elaboração dos instrumentos para a recuperação da informação e o acesso aos documentos. A estrutura organizacional, identificada nos formulários da Metodologia DIRKS possui entidades que têm como objetivo buscar informações em metadados tais como: a estrutura hierárquica, as unidades administrativas responsáveis pela função, atividade e transação, cargo de presidente de comissão permanente de avaliação de documentos, o gerente de projeto e da relação entre as unidades da estrutura organizacional e os projetos, por exemplo.

As entidades dos formulários da metodologia DIRKS, que apontaram para a categoria específica 1.1 **Organização**, mostraram informações relevantes que poderiam ser associadas ao contexto organizacional, tal como apontado por Thomassen (2006). Para o autor, o contexto organizacional compreende a forma como a organização é estruturada internamente, os departamentos, os agentes externos com as quais ela se relaciona e, por fim, um olhar sobre a organização e os fatores que causam impacto no contexto arquivístico.

Ressalta-se que o Princípio da Proveniência fundamenta toda a metodologia DIRKS e aparece explicitado na coleta de dados de todos esses elementos da organização como: relacionamentos administrativos, estrutura hierárquica, nome da unidade, característica e história da instituição, competência, *stakeholder*⁴⁴ e pessoas, como elementos fundamentais para a preservação dos fundos arquivísticos ou para a pesquisa visando sua identificação e, obviamente para a construção dos instrumentos de gestão.

Além dos elementos citados anteriormente, outros foram identificados com o objetivo de buscar informações sobre as possíveis interações administrativas com outras organizações. Da mesma forma, a história da organização e os elementos que caracterizam esta organização são dados importantes para a arquivologia porque revelam informações sobre o contexto de produção dos documentos de arquivo. Estas informações são importantes, também, para fundamentar a elaboração dos instrumentos de gestão. Além disto, a história e os elementos da organização também fazem parte da busca dos subsídios necessários á descrição arquivística.

A categoria específica **1.2 Instrumentos de Apoio** contém os instrumentos necessários para dar suporte ao desenvolvimento das atividades nas organizações e são formados pelos metadados: Padrões da Organização; Requisitos de Arquivamento; Controle Interno e Tecnologia usadas na instituição. Tais metadados compreendem os elementos revelados pela investigação do contexto organizacional que foram levantados para a identificação tanto da estrutura legal quanto dos regulamentos e os padrões responsáveis pelas boas práticas organizacionais, instrumentos que regulam e normatizam as atividades da organização. Esses elementos são, também, apontados tanto pelo CIA, quanto pelo Projeto InterPares como recursos facilitadores para a implantação de uma política de gestão de documentos. Assim,

O CIA (2005) ressalta a importância dos padrões e práticas para a gestão de documentos e a preservação das suas características ao afirmar que “os serviços de arquivo devem facilitar o estabelecimento de políticas, procedimentos, sistemas, normas e práticas concebidas para apoiar os produtores de documentos de arquivo a criar e manter documentos que sejam autênticos, fidedignos e passíveis de ser preservados a longo prazo” (CIA, 2005, p.7).

⁴⁴ *Stakeholder* são pessoas ou instituições envolvidas nas transações ou processos referentes a cada atividade levantadas nos órgãos.

Fonte de informação

A categoria específica **1.3 Fonte de Informação** é formada por: 1.3.1 Documento do Processo Dossiê; 1.3.2 Fonte de Informação Legislação; 1.3.3 Fonte de Informação Bibliográfica e 1.3.4 Fonte de Informação Pessoa e as entidades dos formulários DIRKS. No caso da metodologia DIRKS as fontes de informação são fundamentais, pois revelam não só informações sobre as fontes usadas em cada organização para o desempenho de suas atividades, mas também fornecem os subsídios necessários para a equipe que trabalha na construção dos instrumentos de gestão.

A subcategoria **1.3.1 Documento do Processo Dossiê** representa os elementos do formulário Análise de Atividade - Fluxo da Transação e Identificação de Requisitos de Arquivamento. Esses formulários têm como objetivo o levantamento dos requisitos de arquivamento de cada transação e o registro do seu fluxo, isto é, as ações que geram e recebem documentos e, com isso, permitem a identificação do conjunto documental correspondente a cada dossiê.

Para a gestão de documentos, ter ciência dos documentos que compõem os dossiês existentes na organização é condição fundamental para facilitar o entendimento da relação entre os documentos do mesmo dossiê, os requisitos de arquivamento e a definição dos prazos de guarda. Outro aspecto que merece ser destacado é que, esse conhecimento fornece uma orientação para preservar a ordem original e identificar os conjuntos documentais nos planos de classificação e nos arranjos arquivísticos. Além dos documentos produzidos e recebidos, outras fontes de informação como legislação, materiais bibliográficos e pessoas também foram identificadas durante a implementação do Passo B da metodologia DIRKS. Este levantamento possibilitou confirmar o uso de uma diversidade de fontes de informação para subsidiar a realização das atividades das organizações, assim como identificar e confirmar também, existência de outros tipos de fontes, tais como pessoas, utilizadas nas instituições para a execução das suas atividades. Em todos os casos discutidos, vale ressaltar a importância dessas fontes de informação, tanto para as atividades da gestão de documentos, como para o desenvolvimento de recursos de informação. Ressalta-se que tais recursos são reveladores do contexto organizacional, das necessidades de informação, portadores dos requisitos de arquivamento, da história da organização e essenciais em outras funções arquivísticas como descrição e elaboração dos instrumentos de pesquisa.

Ato Normativo

A categoria específica **1.4 Ato Normativo** representa a diversidade de tipos de atos legais e administrativos, existentes no contexto organizacional. Esse conjunto de atos como leis, decretos, regulamentos e portarias, é resultado do princípio da publicidade dos atos da administração pública que regulam as atividades desenvolvidas nas organizações e, ao mesmo tempo dão à devida transparência às ações da administração pública.

Programas, Projetos

E, por último, a categoria específica **1.5 Programas, Projetos**, dividida em duas subcategorias: Programas e Projetos terceirizados. Os programas e projetos são os meios pelos quais as políticas são implementadas pelas instituições públicas ou privadas. No domínio organizacional analisado, os programas foram identificados durante a realização da atividade de levantamento do contexto organizacional. Nesse caso específico, verificou-se o contexto de negócio, tendo como objetivo conhecer os detalhes do programa, estratégias ou procedimentos de conformidade e a existência ou não de programa de gestão de risco, associados às atividades da organização.

Metadados da Categoria fundamental 2 Energia

A categoria fundamental **2 Energia** compreende as ações representadas no domínio organizacional por método, processo, operação e técnica. Nos formulários da Metodologia DIRKS fazem parte dessa categoria os seguintes elementos: **2.1 Função; 2.2 Atividade; 2.3 Transação e 2.4 Processo**, metadados fundamentais que representam o modelo conceitual e que retratam as funções e atividades desenvolvidas pelas organizações

Para discutir o papel da função resgata-se a sua definição presente na Norma Internacional para Descrição das Funções, (ISDF) em 2007. Nesta norma a função é apresentada como “qualquer objetivo de alto nível, responsabilidade ou tarefa prescrita como atribuição de uma entidade coletiva pela legislação, política ou mandato”. E, ainda mostra que, as funções podem ser “decompostas em conjuntos de operações coordenadas, tais como subfunções, procedimentos operacionais, atividades, tarefas ou transações” (CIA, 2008, p. 13).

Outro aspecto importante atribuído às funções é o seu papel na elaboração dos instrumentos de gestão: plano de classificação e tabela de temporalidade. A construção desses instrumentos pressupõe a identificação das funções de responsabilidade do organismo produtor, bem como as atividades desenvolvidas referente à área-meio e às áreas-fim da organização.

A função, reconhecida na literatura como mais estável que as estruturas organizacionais, pode servir como princípio para o agrupamento de classes em um plano de classificação, devido a sua característica de associação com as atividades do negócio da organização e por representar as unidades menores, a transação e o conjunto documental registradas da classe mais geral até a mais específica.

Observou-se que a aplicação do princípio da proveniência à medida que o documento, fruto da ação geradora de cada transação, representa o conjunto documental produzido como evidências dessas atividades, revelando acerca da sua função o que, por fim resulta na própria definição de arquivos. Esse papel de explicar a proveniência dos documentos pela descrição das funções, também foi apontada pelo CIA (2008, p.13) por facilitar aspectos como produção e uso dos documentos. Além disso, a explicitação da proveniência pode “ajudar a situar os documentos com mais segurança no contexto de sua produção e uso” nas organizações.

Na definição do termo Função a ISDF expande a sua amplitude. Ou seja, foram considerados nesse grande “guarda-chuva” que simboliza a função, as subcategorias que nesse estudo foram assim agrupadas em: **2.2 Atividade e 2.3 Transação**.

Já a categoria específica **2.4 Processo** é formada pelos metadados: 2.4.1 Estratégia; 2.4.2 Terceirização e 2.4.3 Gestão de documentos. Esta categoria específica compreende os fatores que afetam as práticas de arquivamento na instituição, bem como os elementos que confirmam ou não a existência de programas de gestão de documentos. Estes dados foram identificados na primeira etapa da Metodologia DIRKS e foram utilizados para complementar as informações acerca da investigação preliminar, no Passo A.

Metadados da Categoria fundamental 3 Espaço

A categoria fundamental **3 Espaço** abriga os elementos dos formulários que permitem a localização da instituição e de pessoas que, de alguma forma, relacionam-se com ela. As subcategorias que se encaixam categoria espaço são: **3.1 Endereço da Unidade, 3.2 Telefone, 3.4 Endereço do Site**.

As categorias específicas **3.1 Endereço da unidade e 3.2 Telefone** se referem à localização da unidade da organização. O endereço indica a distribuição espacial das unidades da estrutura hierárquica e o telefone complementa as informações acerca do espaço das unidades da organização.

Esses dados possibilitaram, também, estabelecer um canal de comunicação entre o funcionário que trabalha com a gestão de documentos, os funcionários da organização e principalmente com a instituição arquivística. Além disto, os dados referentes á categoria

fundamental 3 Espaço compreendem o endereço do *site* que, abriga a ementa do texto legal, o endereço eletrônico que inclui os *e-mails* da unidade da organização e do gerente do projeto. Neste caso, a identificação do endereço eletrônico significa a existência de um canal de comunicação com a instituição de origem e com a instituição arquivística.

Metadados da Categoria fundamental 4 Tempo

Junto à categoria **4 Tempo**, foram agrupados os conceitos que relacionados ao significado de medida de duração dos fenômenos. Neste sentido, essa categoria aparece como o elemento que registra o tempo de guarda, no instrumento de gestão de documentos utilizado para a atividade de avaliação dos documentos arquivísticos nas instituições. Na Tabela de Temporalidade o tempo é uma variável fundamental, resultado da reflexão sobre necessidade de informação vinculada à necessidade de evidência e o registro de análises acerca de elementos como a função, a atividade, a legislação, os *stakeholders* e os documentos produzidos e recebidos. Com isto, o tempo definido e registrado na tabela de temporalidade é fruto destas análises e é fruto da tomada de decisão, em relação à guarda dos documentos de arquivo.

Vale ainda, ressaltar a importância entre a evidência do uso dos documentos como fonte de consulta para a sociedade e para pesquisadores sob os aspectos históricos, culturais e científicos identificados na norma ISO 15.489. Desta forma, observou-se claramente uma relação entre essas evidências e a necessidade da definição criteriosa de por quanto **tempo** deve ser guardada a documentação. Nessa categoria fundamental foram reunidas as categorias específicas: **4.1 Ciclo de Vida; 4.2 Data do Ato Normativo e 4.3 Data do Evento.**

Outra categoria específica vinculada à **categoria 4 Tempo** é a **4.1 Ciclo de vida**. Esta subcategoria compreende os metadados 4.1.1 Fase de Arquivamento, 4.1.2 Prazo e 4.1.3 Destinação Final. A primeira significa por quanto tempo os documentos são mantidos nas três idades ou fases de arquivamento fazendo referência à teoria das três idades. Compreendem um duplo sentido para o termo tempo, isto é, representam tanto a idade do documento: fase corrente, intermediária e guarda permanente, quanto os prazos de guarda que serão mantidos em cada uma dessas fases.

E a última divisão da categoria **4.1Ciclo de vida** é a **Destinação final**, também, compreende a tomada de decisão em relação ao tempo. A decisão em torno da destinação final retrata duas possibilidades: eliminação ou guarda permanente. Se a escolha for pela eliminação, isto significa que o prazo de guarda expirou e que o documento cumpriu o seu valor primário, ou seja, cumpriu o objetivo para o qual ele foi criado nas organizações. A

segunda alternativa é para a guarda permanente dos documentos. Neste caso, os documentos que serão recolhidos à guarda permanente ou à terceira idade têm um valor secundário ou informativo e, contêm informações que serão utilizadas pela sociedade para fins de prova ou para pesquisas científicas.

Os metadados **Ciclo de vida, Prazos para a guarda ou Eliminação, os Prazos de guarda e a destinação**, denotam tempo e aparecem registrados no instrumento de gestão tabela de temporalidade presente atividade de avaliação de documentos apontados por Bernardes (1998) e pelo Arquivo Nacional (2005).

6 CONSIDERAÇÕES FINAIS

O sistema categorial construído a partir dos formulários da Metodologia DIRKS reflete domínios organizacionais do ponto de vista da arquivologia, Além disso, esse sistema categorial revela os aspectos culturais da comunidade discursiva, visto que a sua elaboração compreendeu a análise de dados presentes nos 8 (oito) formulários usados para a coleta de dados nos órgãos da Administração do Estado de Minas Gerais.

Considerando, então, o sistema categorial proposto observou-se que os objetos de um domínio podem ser representados a partir das categorias fundamentais de Ranganathan Personalidade; Energia; Espaço e Tempo – e categorias específicas identificadas a partir de um novo agrupamento.

Nesse processo, cabe aqui ressaltar a descrição de alguns metadados fundamentais para amparar a criação dos instrumentos de gestão. A metodologia usada nesta pesquisa e as decorrentes reflexões evidenciam os princípios teóricos da área: os princípios do respeito aos fundos, da ordem original e a abordagem da teoria das três idades que fundamentaram a interpretação dos elementos dos formulários de coleta de dados. Da mesma forma, os levantamentos do contexto organizacional, especificamente do ambiente regulatório, serviram como subsídios importantes para fornecer os componentes das funções e atividades da organização e, permitiram reunir os elementos do contexto de geração e uso dos documentos de arquivo, bem como para nortear a criação dos instrumentos de gestão.

No delineamento do domínio organizacional ressaltam-se os objetos extraídos do contexto funcional, de acordo com Thomassem (2006), oriundos da missão da organização e representados pelas funções, atividades, transações e os processos estão identificados como elementos fundamentais nos formulários Levantamento de Atividades de Negócio.

Cumpré finalmente ressaltar que os campos presentes nos formulários DIRKS e que foram analisados não representam os dados e sim os metadados, destinados à modelagem de

um domínio organizacional e a construção dos instrumentos de gestão de documentos de arquivos. Sheperd & West (2003), em estudo sobre os metadados, a partir da norma ISO 15.489 e da norma ISAD45, definem os metadados para a gestão de documentos. De acordo com as autoras, esses recursos referem-se aos dados necessários para gerenciar, recuperar e acessar os documentos de arquivo, ao longo da sua vida. Metadados é a “informação estruturada ou semi-estruturada que permite a criação, a gestão e o uso de documentos através do tempo e do domínio no qual eles foram criados”, (Wallace, 2001 citado por Sheperd & West, 2003, p.11). Nesse sentido, pode-se afirmar que há indícios que o sistema categorial decorrente da presente pesquisa possa ser visto como um padrão de metadados que podem ser usados como orientadores para o processo de construção dos instrumentos de gestão fazendo uso dos três primeiros passos a Metodologia DIRKS.

REFERÊNCIAS

ARQUIVO NACIONAL (Brasil). **Dicionário brasileiro de terminologia arquivística**. Rio de Janeiro: Arquivo Nacional, 2005. 232p. Disponível na internet: <http://www.conarq.arquivonacional.gov.br/Media/publicacoes/dicionrio_de_terminologia_arquivstica.pdf>. Acessado em: 30/04/2014.

AUSTRALIA. NATIONAL ARCHIVES OF AUSTRALIA. **Designing and Implementing Recordkeeping Systems (DIRKS) Manual**. Canberra: National Archives of Australia, 2007. disponível em: <<http://www.naa.gov.au/records-management/publications/DIRKS-manual.aspx>>. Acesso em: 22 nov. 2007.

BERNARDES, Ieda Pimenta. **Como avaliar documentos de arquivo**. São Paulo: Arquivo do Estado, 1998. (Projeto como fazer, v.1)

BRASCHER, Marisa; CAFÉ, Lígia. Organização da Informação ou Organização do Conhecimento? In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB), 9., 2008, São Paulo. **Anais...** São Paulo: USP, 2008.

CAMPOS, Maria Luiza de Almeida; GOMES, Hagar Espanha. Organização de domínios de conhecimento e os princípios ranganathianos. **Perspect. cienc. inf.**, Belo Horizonte, v. 8, n. 2, p. 150-163, jul./dez. 2003.

CIA - Conselho Internacional de Arquivos. Comité de Arquivos Correntes em Ambiente Electrónico. **Documentos de arquivo electrónicos: manual para arquivistas**. Paris: CIA, 2005. 74p. (ICA, Estudo n.º 16 – Manual para Arquivos Electrónicos). Disponível em: <www.ica.org>. Acesso em: 12 set. 2009.

CIA - Conselho Internacional de Arquivos. **ISDF: Norma internacional para descrição de funções**. Rio de Janeiro: Arquivo Nacional, 2008.76 p.

45 ISAD – Norma Geral Internacional de Descrição Arquivística. Esta norma estabelece diretrizes gerais para a preparação de descrições arquivísticas. Deve ser usada em conjunção com as normas nacionais existentes ou como base para a sua criação.

Disponível em: <<http://pt.slideshare.net/doritchka/brascher-e-caf-organizacao-da-informao-ou-do-conhecimento>>. Acessado em 10/08/2012

GARSHOL, Lars Marius **Metadata? Thesauri? Taxonomies? Topic Maps! Making sense of it all**. Ontopia, Oslo, 2004. <http://www.ontopia.net/>Disponível na internet: <<http://www.ontopia.net/topicmaps/materials/tm-vs-thesauri.html>>. Acessado em 05/11/06

ISO 15489. Information and documentation -- Records management -- Part 1: General. Disponível em: <http://www.iso.org/iso/catalogue_detail?csnumber=31908>. Acessado em 05/06/14.

LOURENÇO, Cíntia de Azevedo Lourenço. **Modelagem de dados como ferramenta de análise de padrões de metadados em bibliotecas digitais**: o padrão de metadados brasileiro para teses e dissertações segundo o modelo entidade-relacionamento. 2005. 161f. Tese (Doutorado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2005.

MIRANDA, M. L. C. A organização do conhecimento e seus paradigmas científicos: algumas questões epistemológicas. **Informare**: Cadernos do Programa de Pós-Graduação em Ciência da Informação, Rio de Janeiro, v. 5, n. 2, p. 64-77, 1999.

RONDINELLI, Rosely Curi. **Gerenciamento arquivístico de documentos eletrônicos**: uma abordagem teórica da diplomática arquivística contemporânea. 4. ed. Rio de Janeiro: FG, 2005. 158p.

SHEPERD, Elizabeth; WEST, Victoria. Are ISO 15.489-1:2001 and ISAD(G) compatible? Part 1. **Records Management Journal**, v.13, n.1, p.9-23, 2003

THOMASSEN, Theo. Uma primeira introdução à arquivologia. **Arq. & Adm.**, Rio de Janeiro, v. 5, n. 1, p. 5-16, jan./jun. 2006.

A ORGANIZAÇÃO DO CONHECIMENTO ARQUIVÍSTICO: O PRINCÍPIO DA PROVENIÊNCIA E O SISTEMISMO

THE ORGANIZATION OF ARCHIVAL KNOWLEDGE: THE PRINCIPLE OF PROVENANCE AND THE SYSTEMISM

Josemar Henrique de Melo
Américo Augusto Nogueira Vieira
Guilherme Ataíde Dias

Resumo: Analisa quais as alterações ocorrem no processo de organização dos acervos arquivísticos a partir do incremento da perspectiva sistêmica junto ao Princípio da Proveniência, que rege atualmente os procedimentos da Arquivologia. O texto compreende quatro subseções: uma introdução que abarca questões de ordem lógico-metodológicas com vistas a possibilitar um olhar sobre a Arquivística inserindo-a no contexto das ciências. Após, faz-se um recorte histórico-conceitual do surgimento do Princípio da Proveniência e, simultaneamente, inserindo a perspectiva sistêmica como um mecanismo crítico que permite um olhar mais abrangente. A pesquisa para realização do presente artigo partiu de ampla base bibliográfica e documental e foi desenvolvida em colaboração por pesquisadores de diversas universidades com vistas à organização do conhecimento, na forma sistêmica, em Arquivologia. Ao final, conclui-se que a cientificidade da Arquivologia passa pela revisão dos seus enunciados teóricos, sendo o sistemismo, como forma de organização da informação em arquivística, uma possibilidade plausível de efetivar as inter-relações necessárias com outras áreas de conhecimento e que não há, em regra, nenhuma contradição aparente entre o Princípio da Proveniência e a inclusão da perspectiva sistêmica em Arquivologia.

Palavras-Chave: Princípio de Proveniência. Arquivologia. Sistema. Problema da Demarcação.

Abstract: Analyzes what changes occur in the organization process of archival collections from the systemic perspective increase together with the Principle of Provenance, which currently governs the Archival procedures. Four sub-sections compose the text: an introduction that covers issues on logical and methodological order with a view that enables a look at Archivology inserting it in the context of science. Afterwards, it is done a historical-conceptual cut on the emergence of the Principle of Provenance and simultaneously inserting the systemic perspective as a critical mechanism that allows a more comprehensive understanding. The research steps for writing this paper was grounded on an extensive literature and documental base and was developed in collaboration with researchers from several universities aiming to the systemic organization of knowledge in Archivology. It is concluded that the scientific aspect of the Archivology passes by the review of its theoretical statements. A Systemic perspective as a way of organizing information in Archivology is a plausible possibility of establishing the necessary interrelationships with other areas of knowledge and that there are no apparent contradictions between the Principle of Provenance and the inclusion of the systemic perspective in Archivology.

Keywords: Principle of Provenance. Archival Science. System. Problem of Demarcation.

1 INTRODUÇÃO

Uma das questões fundamentais enfrentadas por um novo campo de conhecimento é a avaliação de se saber se está, o referido campo, destinado a ser uma ciência (tal como a

Física)⁴⁶, ou uma ciência e, ao mesmo tempo, uma técnica (que é o caso do Direito, da Medicina e também das várias engenharias), ou se é apenas um conjunto de técnicas (como é o caso de Comércio Exterior, Radialismo, Gastronomia, Produção Cultural, etc.) ou mesmo se esta se encontra no campo da Arte (como é o caso da pintura, que apesar de possuir sistematicidade e até mesmo ser passível de investigação não é uma ciência)⁴⁷. Não ser ciência não implica em não ser importante para a condição do humano; é apenas uma classificação associada a certos usos e critérios. A Música e a Gastronomia, por exemplo, mesmo não sendo ciências, fazem da vida humana algo mais sublime.

A avaliação sobre um campo de conhecimento ser ciência (ou não) deve cumprir minimamente o que se denomina de *critério da demarcação*. Karl Popper colocou o *Problema da Demarcação* como o problema de estabelecer um conjunto de critérios hábeis para distinguir “as ciências empíricas, de uma parte, e a Matemática e a Lógica, bem como os sistemas metafísicos, de outra.” (POPPER, 1993, p.35). Ainda na *Lógica da Pesquisa Científica*, Popper (1993, p.31) enfatiza que a tarefa dos pesquisadores de uma dada ciência é construir teorias e colocar, então, as teses dessa teoria à prova; isto é, que as teses possam ser sujeitas, por *critérios de testabilidade*, ao que denominou de “possibilidade de falsificação” ou sujeitas à “falseabilidade”. Teses que não sejam sujeitas à falseabilidade (aquelas que são sempre verdadeiras, por exemplo, às subjunções ou bijunções de caráter tautológico) não poderiam caracterizar um campo de conhecimento como científico.

Segundo Vieira (2004, p.20): “Diz-se que um enunciado molecular é uma tautologia se, depois de verificadas todas as possíveis valorizações de seus enunciados primos, o enunciado molecular for verdadeiro em todos os casos”. Nesse sentido, de ser um enunciado uma tautologia, a única testabilidade possível é a lógica (não é empírica, como nos quer indicar Popper em seu “critério de falseabilidade” apostado na *Lógica da Pesquisa Científica*). Se o enunciado molecular em teste for de Lógica ou de Matemática, então tais campos do saber ainda serão aceitos pelo critério de cientificidade por serem sistemas de representação científicos (que Popper prefere denominar ciências não empíricas). Por outro lado, se os enunciados moleculares desse campo forem tautologias (e não se encaixarem como gramáticas, “linguagens” ou sistemas de representação, tais como a Lógica ou a Matemática)⁴⁸ ou se, seus enunciados-tese não sejam passíveis de critérios de falsificação, isto

⁴⁶ Que é entendida como a “ciência referência” ou mesmo “ciência mãe”.

⁴⁷ Ainda que a História da Pintura possa ser um subcampo da ciência da História.

⁴⁸ Quine (1972) foi o primeiro a perceber que se poderia perceber a Lógica como uma espécie de gramática.

é, quando sempre haveria uma *fuga da falseabilidade* do teste empírico por solução *ad hoc*⁴⁹ da teoria ou previsão em teste, então tal campo de saber não seria científico.

Portanto, se aceitas as considerações de Popper (1933, p.31) que a tarefa do cientista é contribuir para o desenvolvimento de teorias, então devemos ter em mente que as teorias necessitam de **princípios**, conceitos, axiomas, mecanismos de argumentação e finalmente de “enunciados teses” (ou previsões) sujeitos a falseabilidade.

Nesse sentido, Stiglitz e Walsh (2003, p.17) nos apontam: “As previsões de uma teoria têm a forma: se aumentarmos um imposto num mercado concorrencial, então a produção diminuirá e os preços aumentarão”; portanto, uma subjunção, tal como antes os já havíamos apontado⁵⁰, além do fato que é muito clara a verificação da **possibilidade de falseabilidade**, tanto do antecedente/hipótese quanto da conclusão que se segue desta, quanto também da subjunção (tese ou previsão como denominam Stiglitz e Walsh na citação) como um todo.

Ressalte-se que este artigo foi desenvolvido a partir da investigação do problema de compatibilização do Princípio de Proveniência com propostas de inclusão do sistemismo em teorias arquivísticas; tais como, lançadas pelo Professor Armando Malheiro da Silva e seus colaboradores de pesquisa (SILVA, et. al., 1999).

Assim, o presente trabalho traz a lume, tanto de forma sistemática, como através de uma breve revisão histórica, o desabrochar de um dos princípios mais utilizados na Ciência Arquivística: o Princípio da Proveniência; ou, como também é conhecido o Princípio de Respeito aos Fundos em Arquivística. Tal esforço de pesquisa, em se perseguir os fundamentos teóricos da Arquivística fica plenamente justificado ao cotejarmos tal esforço e objetivo de pesquisa com as teses de Popper anteriormente elencadas. A pesquisa para realização do presente artigo teve ampla base bibliográfica e documental, mas focou-se na parte mais estritamente de arquivística principalmente nos textos de Michel Duchein, sendo desenvolvida em processo colaborativo por pesquisadores de diferentes universidades brasileiras.

2 O PRINCÍPIO DE PROVENIÊNCIA COMO UMA PERSPECTIVA CUSTODIAL

⁴⁹ “Solução ad hoc” é uma adição de hipótese(s) ou premissa(s) estranha(s) a uma teoria (incorporadas essas à última hora) para tentar salvar a teoria de ser demonstrada como falsa.

⁵⁰ Havíamos apontado que os enunciados tese (ou *previsões* como denominam Stiglitz e Walsh) poderiam ser subjunções (que muitos conhecem pela denominação de implicação) ou bijunções, sendo esta última um enunciado que é subjunção nos dois sentidos (forma tipo “p ↔ q”).

Historicamente o suporte sistematizador da Arquivística custodial surge durante o processo revolucionário francês iniciado no século XVIII⁵¹. A Revolução Francesa causou não só grandes modificações no campo político mundial como também mudaram os padrões da administração tradicional o que, conseqüentemente, provocaram enormes impactos no que toca aos arquivos com o estabelecimento do chamado “Princípio de Procedência” ou de “Princípio de Respeito aos Fundos”⁵² estruturado por Natalis de Wailly, chefe da secção administrativa dos arquivos departamentais do Ministério do Interior de França.

Nos anos que se seguem após a Revolução, os governantes franceses sentiram “que era preciso criar um serviço novo, que se encarregasse de zelar pelos documentos oficiais em que passou a assentar o regime” (SILVA, *et al.*, 1999, p.100). Desta forma, neste período de turbulência fundaram em 1794 o *Archive National* da França para onde migraram acervos de diversas procedências do Antigo Regime sem nenhuma ordenação lógica desses documentos, provocando por este meio à quebra das suas unidades sistêmicas⁵³. Michel Duchein (1989) esclarece-nos ainda que:

os dois primeiros diretores desse grande depósito, Armand Camus e Pierre Daunou, conceberam o projeto de tratar essa massa de arquivos como um único conjunto documentário, dividido em cinco ‘secções’ cronológico-metódicas: secção legislativa para os papéis das Assembleias revolucionárias; secção administrativa para os papéis dos novos ministérios; secção de domínio para os títulos de propriedade do Estado; secção judicial para os documentos dos tribunais e por fim a secção histórica. Esta última era formada por documentos considerados (de maneira arbitrária) como particularmente interessantes do ponto de vista da história e extraído de outras secções. Estes documentos foram classificados (ou pretendidos classificados) por locais, datas, reinos, etc. de tal forma que em muitos casos a origem dos documentos foi impossível de encontrar tudo estando inextricavelmente misturado e disperso (DUCHEIN, 1986, p.15).

Em meio ao caos produzido por esta enxurrada de documentos, foi então proposto por Natalis de Wailly em uma circular de nº 14 datada em 24 de Abril de 1841 intitulada “*instructions pour la mise en ordre et le classement des archives départementales et*

⁵¹ O “processo revolucionário” está aqui sendo entendido como aquele que tem como marco inicial a “Queda da Bastilha” e que se estende até, e inclusive, o “Período do Terror” (agosto de 1792 até 27 de julho de 1794).

⁵² É importante destacar que o Princípio de Procedência, como apresenta Martín-Pozuelo Campillos (1996) não foi resultado de um único momento, nem produzido num único local, pois já havia sido proposto e utilizado em outros países e em outros períodos, antes da famosa circular de Natalis de Wally.

⁵³ Se, por um lado, a Revolução Francesa criou uma espécie de caos ao reunir e misturar os documentos de diferentes entidades, por outro lado, os revolucionários decidiram abrir o acesso público dos acervos a todos os cidadãos, dando, desta forma, o primeiro passo ao que poderíamos alegoricamente denominar de “democracia informacional.”

*communales*⁵⁴,” o que ficou conhecido como “Princípio da Proveniência” (ROUSSEAU; COUTURE 1998, p.80) o que Duchein (1986, p.16) definiu como “a certidão de nascimento da noção de fundos de arquivo”. Em seu aspecto fundamental esta regra da Arquivística, como nos alertam Jean-Yves Rousseau e Carol Couture (1998, p.80), atribui uma obrigatoriedade em: “agregar os documentos por fundos, isto é, reunir todos os títulos (documentos) provenientes de um corpo, de um estabelecimento, de uma família ou de um indivíduo, e dispor segundo uma determinada ordem os diferentes fundos”. Este primeiro princípio, que se pode denominar “Princípio de Proveniência”, ou também como “Princípio do Respeito aos Fundos em Arquivística”, vem aludir à integridade externa dos acervos, principalmente no que concerne a separação entre unidades documentais de outras diferentes entidades. Contudo, no que respeita a ordem interna destas mesmas unidades, a circular proposta por Natalis de Wailly faz, inicialmente, uma separação dos documentos anteriores e posteriores a 1790, sendo à disposição dos acervos diferente para os dois períodos, o que toma como base as diversas “espécies de arquivo”.

Durante o século XX, o acatamento do Princípio de Proveniência foi sendo consolidado através de pesquisadores que têm apresentado trabalhos postulando melhorias na aplicação da proposta de Natalis de Wailly; entre eles salientamos o trabalho de Michel Duchein, que iremos em seguida analisar foi publicado originalmente em 1977 com o título *Le respect des fonds em Archivistique: principes théoriques et problèmes pratiques*⁵⁵. Este texto foi utilizado por diversos autores⁵⁶ para contextualizar a aplicação do Princípio dos Fundos nas organizações e fundamentar teoricamente a disciplina Arquivística. No referido texto Duchein pontua a complexidade que as organizações podem ser e ter, e também as dificuldades em colocar em prática, da melhor maneira possível, o Princípio de Respeito aos Fundos em Arquivística.

O Princípio de Proveniência foi enunciado e aplicado, tendo em vista solucionar o caos do Arquivo Nacional francês⁵⁷; porém sua real dimensão e sua aplicabilidade, associadas

⁵⁴ A aplicação do Princípio de Proveniência já era utilizada em alguns arquivos pela Europa de maneira empírica antes da sua enunciação por Natalis de Wailly, ou mesmo antes da aceitação geral deste como um dos princípios norteadores da Arquivística.

⁵⁵ Este texto foi traduzido e publicado no Brasil pela revista da Associação dos Arquivistas Brasileiros em 1986 com o título: *O respeito de fundos em Arquivística: princípios teóricos e problemas práticos*.

⁵⁶ Entre estes autores podemos destacar: Belloto (1991), Rousseau; Couture (1998), Heredia Herrera (1993), Albuquerque; Souto (2013), entre outros.

⁵⁷ Naquela oportunidade, a aplicação do referido princípio assumido se deu de maneira estanque; isto é, não foi aplicada em toda sua gênese, que vai do respeito do princípio desde sua entidade

à dinâmica organizacional das instituições que forma seus arquivos, não tinham ainda sido testadas⁵⁸. A dinâmica das instituições é apresentada e a formação dos fundos arquivísticos é então problematizada a partir de cinco questionamentos levantados por Duchein.

Neste texto, o Duchein aponta que (1986, p.14) “o respeito aos fundos vem sendo considerado, desde a segunda metade do século XIX, como o princípio fundamental da Arquivística”, por ter sido este preceito que conseguiu constatar a necessidade de manter os documentos de arquivo unidos organicamente e que é a sua própria razão de ser, dando-lhe assim o estatuto de base teórico-prática da Arquivística. Para ele, qualquer trabalho arquivístico fora deste âmbito é arbitrário, subjetivo e sem rigor, correndo-se o risco de não saber onde encontrar um documento (DUCHEIN, 1986, p.16).

O referido princípio fortalece a ideia de individualização dos arquivos produzidos pelas mais diversas instituições, tendo em vista que o documento de arquivo funciona como parte de um conjunto e sua existência só se justifica, compreendendo com exatidão quem o produziu, sem misturas de documentos de outras proveniências. Assim, para o arquivista e demais estudiosos, a correta organização e integridade (externa e interna) de um arquivo deve representar a maneira lógica do funcionamento dos organismos que produziram os documentos e suas inter-relações, facilitando o entendimento das funções e atividades pertinentes a estas instituições dentro de uma contextualidade própria.

Na terceira parte do texto Duchein apresenta suas propostas para melhor adaptar a realidade encontrada nas instituições e a proveniência dos acervos, tendo em consideração que havia inúmeros problemas impostos pelas dificuldades nas *praxis* resultantes da aplicação da teoria dos fundos ou dos princípios de proveniência e da ordem interna, incluindo-se também o que o autor pontua de “ambiguidade do próprio termo”. Tais problemas são definidos em cinco pontos essenciais que veremos na sequência.

O primeiro problema seria “a definição de fundo em relação à hierarquia dos organismos produtores”. Isto é, este problema lançado por Duchein diz respeito à complexidade dos vínculos de subordinação entre instituições (são as ocorrências de subdivisões ou formação de filiais) e a estruturação dos fundos arquivísticos. Neste caso, são

produtora até sua destinação final. Ali o foi somente para separar o acervo “confuso” que já se encontrava produzido.

⁵⁸ A aplicabilidade do Princípio de Proveniência passa a ser verificado a partir do momento em que a ideia do que vem a ser arquivo amplia-se, ou seja, quando os documentos que formam o arquivo são produzidos e não apenas quando, perdido o interesse a instituição produtora, sendo transferidos para um setor de arquivo. Sobre este ponto ver o texto *Arquivística general* de Heredia Herrera (1993, p.98).

os arquivos das subdivisões tomados como fundos distintos ou pertencentes aos fundos das instituições a qual estão vinculados? Este é um problema prático enfrentado pelos arquivistas das mais variadas instituições, sejam elas públicas ou privadas, e podem-se multiplicar situações no domínio complexo das hierarquias.

Para responder a esta questão Duchein propõe a possibilidade de duas diferentes soluções: maximalista ou minimalista. A primeira define um fundo único a partir da maior unidade orgânica, do qual todos estão vinculados, mas não formam fundos arquivísticos. A segunda possibilidade, ao contrário, permite que as unidades orgânicas vinculadas possam formar distintos fundos arquivísticos. Porém, para identificar esta capacidade, as organizações devem observar:

- i) possuir uma existência jurídica resultante de lei, decreto, resolução, etc.;
- ii) ter atribuições precisas definidas por norma;
- iii) ter subordinação conhecida firmada por lei;
- iv) ter uma chefia com poder de decisão no seu respectivo nível hierárquico;
- v) ter uma organização interna fixada.

Duchein destaca o cuidado que se deve ter ao escolher uma ou outra das soluções propostas, pois podem incorrer na formação de fundos documentais muito extensos (maximalista) ou uma pulverização muito grande que obrigaria a criação de seções de arquivos (minimalista). Além disso, o autor nos alerta para um elemento de extrema importância (DUCHEIN, 1986, p.21): “parece-nos inevitável introduzir na Arquivística uma nova noção: a de uma hierarquia de fundos correspondendo à hierarquia dos organismos, acarretando a subordinação de fundos em relação a outros.” Este aspecto demonstra a necessidade de inter-relação entre fundos que não foi pontuada na formulação do princípio.

O segundo problema versa sobre as variações de competência dos organismos produtores de arquivo e suas repercussões no plano arquivístico. Essas variações devem ser cotejadas regularmente pelos profissionais de arquivo, pois estas instituições modificam-se no decorrer do tempo, sendo que tal problema foi analisado por Duchein (1986) da seguinte forma:

- i) casos de supressão de competências;
- ii) casos de criação de competências;
- iii) casos de transferências de competências e;
- iv) casos de competências temporárias.

As variações de competências são ocorrências comuns tanto em instituições públicas como organizações privadas para se adequar a realidade apresentada: criação de setor(es)

extinção de divisões ou transferência de atividades fazem parte do dia-a-dia administrativo das instituições. De uma forma geral, os casos (i), (ii) e (iv) não afetam a identificação de um fundo arquivístico, o próprio autor reconhece que este é “um falso problema, (...) a natureza do fundo não se modifica” (DUCHEIN, 1986, p.22). Neste sentido, apenas as séries documentais dos fundos arquivísticos são ampliadas ou extintas, permanecendo no próprio fundo. Mas, no terceiro caso é necessário entender todo o percurso das instituições para saber se houve alterações orgânicas, quais foram em que nível se deu e com elas as transferências de séries documentais para outros órgãos.

O terceiro problema seria sobre (DUCHEIN, 1986, p.23) “(...) o que entendemos, exatamente, por proveniência de um fundo?” A questão pode inicialmente até parecer redundante, porém é válida quando inclui a transferência do acervo para outra instituição e mais tarde o recolhimento para o Arquivo Público em dois tipos diferentes: (i) agrupados, mantendo a individualidade ou (ii) misturados ao dos organismos que o receberam. Neste aspecto o autor destaca as dificuldades de cunho administrativo tais como: mudança de nome, estrutura e competência do organismo produtor ou erros teóricos e práticos quando o recolhimento para as instituições de Arquivos Públicos não percebem ou não compreendem os diferentes fundos existentes e dispersos sem identificação dentro daquele acervo.

Para este caso o Duchein propõe como solução a manutenção da integridade do fundo arquivístico, excetuando-se os casos em que os acervos perderam completamente sua identidade quando foram transferidos, para estes a procedência passa a ser o do recolhimento e não o da produção.

O quarto problema seria sobre a classificação “fundos abertos e fundos fechados”. A grande reflexão para esta questão baseia-se na real extinção do organismo ou apenas na alteração de nome ou funções. Encontra-se intrínseco a este problema a discussão que durante algum tempo existia na Arquivologia: considerar como arquivo apenas aqueles em que os fundos estejam completos, ou seja, encerrados. Tal discussão já está superada, pois (DUCHEIN, 1986, p.25) “(...) todo documento gerado por um organismo e por ele conservado pertence, desde a sua criação, ao fundo de arquivo deste organismo”. Entretanto, este aspecto dificulta a definição de fundos fechados quando não se tem clareza na continuação ou extinção dos organismos quando ocorre a transferências de todas as suas competências para outras entidades criadas com as mesmas funções, porém com outro nome. Nestes casos não existe encerramento dos fundos, apenas alterações administrativas que não impedem a continuidade do fundo Arquivístico, pois a nova instituição continuará exercendo as mesmas funções. Porém, se houver transferência de um fundo de uma instituição extinta para outra

instituição já existente, o fundo transferido deve permanecer encerrado e distinto dentro do organismo que o recebeu.

O último e quinto problema seria sobre se (DUCHEIN, 1986, p.26): “(...) o respeito aos fundos implica na obediência à sua classificação interna de origem ou não”? Esta última questão apresentada por Duchein refere-se principalmente a ordem interna. Aspecto que não foi pontuado na criação do Princípio de Proveniência e sim no Manual para Classificação e Descrição dos Arquivos também conhecido como *Manual dos Arquivistas Holandeses*⁵⁹. Este segundo princípio arquivístico (ordem interna) tem sido visto por alguns autores da área como um desdobramento do Princípio de Respeito aos Fundos e destaca a necessidade de manutenção da ordem primitiva dada pela entidade produtora ou receptora dos documentos. Os dois princípios iniciam o que se denomina de “fase técnica e custodial” da Arquivologia como destacam Silva e outros (SILVA, et al., 1999).

Vale salientar que, mesmo com uma explicação sobre a terminologia por ele utilizada, ocorre durante toda a apresentação deste último problema uma confusão na utilização dos termos “classificar” e “ordenar” (DUCHEIN, 1986, p.28), “(...) se o organismo adota uma classificação para os seus arquivos correntes (seja ela alfabética, metódica, etc.)”, tratando os dois termos como o mesmo significado, quando na verdade eles tratam de atividades sequenciais, mas distintas como pontua Renato Tarciso Barbosa de Sousa (2007):

é necessário estabelecer o significado e uso de importantes conceitos relacionados ao processo classificatório. São eles: classificação, ordenação, arquivamento, codificação e instrumento de classificação. (...) utilizaremos o termo ‘classificação’ para identificar a ação intelectual de construir esquemas para agrupar os documentos a partir de princípios estabelecidos. A ‘ordenação’ como forma de disposição dos tipos documentais dentro das divisões estabelecidas no esquema de classificação. O ‘arquivamento’ como a ação física de colocar os documentos em pastas ou caixas orientadas pelo esquema de classificação e pela ordenação definida (SOUSA, 2007, p.85).

Desta forma, o autor questiona se o respeito dos fundos implicaria dizer também a manutenção pela classificação interna de origem em dois casos:

- (i) organismos de estruturas simples;
- (ii) organismos de estruturas complexas.

⁵⁹ Esta obra publicada em 1898 foi de grande importância para a comunidade arquivística por tratar de problemas fundamentais para esta área do conhecimento, principalmente no que toca a organização interna dos fundos. Este livro teve grande alcance por ter sido traduzido para diversos idiomas, tornando-se “marco na evolução da Arquivística, por ter aberto uma nova era para a afirmação da disciplina [...] marca o predomínio da vertente teórica” (SILVA, et al., 1999, p.115).

Tanto num caso como no outro Duchein pontua ser bem-vindo o Princípio de Ordem Interna, porém, destaca que a classificação/ordenação (confusão latente nesta parte do texto) deve privilegiar os aspectos arquivísticos (compatíveis com o Princípio do Respeito aos Fundos), que não prejudique a pesquisa, que tenha sido posta em prática e que seja conservada por um longo período. Portanto, se não for possível seguir estes critérios, então (DUCHEIN, 1986, p.30): “devemos renunciar à classificação dada pelo organismo produtor e proceder como se o fundo não tivesse recebido qualquer classificação antes de chegar às mãos do arquivista”.

3 A INCLUSÃO DA PERSPECTIVA SISTÊMICA

As transformações provocadas pelo incremento das tecnologias de informação aos arquivos foram fundamentais para iniciar as discussões sobre o objeto da Arquivologia, pois (SCHMIDT, 2012, p.21): “para alguns teóricos da área, a informação se separa do suporte. E isso para os adeptos da ‘nova’ Arquivologia é um triunfo (...). A partir de então, consideramos importante pensar sobre o estatuto científico da Arquivologia”.

Utilizando-se das ideias de Kuhn (1990) acerca das mudanças de paradigmas autores como Cook (2001) acreditam que as transformações provocadas pelas tecnologias de informação e comunicação causaram profundas modificações no saber e fazer da Arquivologia. Sugere este autor que os arquivistas devam está se posicionando frente às novas demandas, adaptando-se “para ser capaz de pesquisar, reconhecer e articular todas estas radicais mudanças na sociedade e então lidar conceitualmente com seus impactos na teoria arquivística, metodologia e prática” (COOK, 2001, p.17, tradução nossa)⁶⁰.

O debate sobre a cientificidade, ou não, da Arquivologia ainda subsiste⁶¹ (ACEVES, 2011) corporificando-se em duas correntes básicas: (i) entendendo-a como apenas técnica e, por outro lado, (ii) a que assevera ter a Arquivologia uma personalidade científica própria assentes em diferenciados princípios teóricos e metodológicos. E sobre este debate, na busca de bases teórica(s) e metodológica(s) para caracterizar esta área do conhecimento como

⁶⁰ Texto original de Terry Cook (2001, p.17): “to be able to research, recognize and articulate all these radical changes in society and then deal conceptually with their impact on archival theory, methodology and practice”.

⁶¹ Schmidt (2012, p. 24) apresenta, de maneira didática, as perspectivas defendidas e os autores que as defende, sem, entretanto, entrar nas diferenças que existem entre eles. Aqui exporemos apenas os principais, sendo: Arquivologia como técnica: Manuel Romero Tallafigo. Como disciplina técnica: Aurélio Tanodi. Como disciplina em desenvolvimento Theo Thomassen. Como disciplina científica: Heloísa Liberalli Bellotto, Terry Cook, Carol Couture. Como ciência: Antonia Heredia Herrera, Armando Malheiro da Silva e Fernanda Ribeiro.

científica destacamos os trabalhos dos professores Armando Malheiro da Silva, Fernanda Ribeiro, Júlio Ramos e Manuel Luís Leal (1999) com o livro *Arquivística: teoria e prática de uma Ciência da Informação* e, novamente, Malheiro e Ribeiro (2002) em *Das Ciências Documentais à Ciência da Informação*. Destes, destaca-se que (SILVA, et. al., 1999, p.18): “a vastidão de publicações vinda a lume, partindo do equívoco tão vulgarizado que é assumir a técnica e a prática Arquivística como teoria”. Para eles a Arquivística deveria constituir-se a partir de um objeto e um método, definidos com clareza, dentro de parâmetros de uma cientificidade a qual, até então, não possuía, por estar a Arquivística ligada somente a princípios técnicos.

A partir da análise deste desdobramento histórico, Silva e outros (1999) consideram como melhor definição para os arquivos um conceito que tivesse como lastro teórico uma perspectiva sistêmica. Inicialmente cotejada por Bertalanffy (1977), a abordagem, ou perspectiva sistêmica, considera que as atividades humanas podem constituir conjuntos de elementos que interagem ordenada e organicamente, ou seja, formam sistemas, onde:

o todo é mais que a soma das partes, consiste simplesmente em que as características constitutivas não são explicáveis a partir das características das partes isoladas. As características do complexo, portanto, comparadas às dos elementos parecem novas ou emergentes. Se conhecermos o total das partes contidas em um sistema e as relações entre elas, o comportamento do sistema pode ser derivado do comportamento das partes (BERTALANFFY, 1977, p.83).

Assim, a partir da inclusão desta perspectiva, buscou-se reformular as principais noções para esta disciplina e sugeriu-se reescrever o conceito de arquivo como:

Arquivo é um sistema (semi-)fechado de informação social materializada em qualquer tipo de suporte, configurado por dois factores essenciais – a natureza orgânica (estrutura) e a natureza funcional (serviço/uso) – a que se associa um terceiro – a memória – imbricado nos anteriores (SILVA, et al., 1999, p.214):

E a Arquivística é definida como:

uma ciência de informação social, que estuda os arquivos sistemas de informação (semi-)fechados, que na sua estruturação interna e na sua dinâmica própria, quer na interação com os outros sistemas correlativos que coexistem no contexto envolvente (SILVA, et al., 1999, p.214).

Na fundamentação/construção de toda ciência é determinante, para além de uma postura teórica frente ao(s) objeto(s) (a já referida preocupação em formular teorias, tal como apontamos em nossa introdução através de Stiglitz e Walsh), existir também um conjunto de métodos capazes de dialogar com este(s) objeto(s), produzindo uma articulação, ou colagem, entre este(s) e o discurso científico sobre os mesmos. Neste sentido, reforçamos, Silva (et al., 1999) aponta que o Método Hipotético-Dedutivo seria o método a ser utilizado pela

Arquivística. A operacionalidade deste dispositivo se faz através de quatro polos articulados entre si, e que são elementos essenciais para uma investigação científica possuindo as seguintes denominações: o polo epistemológico, o polo teórico, o polo técnico e o polo morfológico.

Assim, a partir da inclusão da perspectiva sistêmica e do método quadripolar, Silva, e outros (1999, p.209) entendem terem trocado os mecanismos tradicionais da Arquivística por um “sistema dinâmico cognoscível através de um dispositivo metodológico integrado”. Por conseguinte, definiram novos princípios como a seguir:

Princípio da acção estruturante – todo arquivo resulta de um acto fundador, individual ou colectivo, formal ou informal, que molda a estrutura organizacional e a sua especificidade funcional em evolução dinâmica.

Princípio da integração dinâmica – todo arquivo integra e é integrado pela dinâmica do universo sistêmico que envolve (interligação e relação com outros sistemas conexos).

Princípio da grandeza relativa – todo arquivo se desenvolve como estrutura orgânica simples (unicelular) ou complexa (pluricelular).

Princípio da pertinência – todo o arquivo disponibiliza informação que pode ser recuperada segundo a pertinência da estrutura organizacional (SILVA *et al.*, 1999, p.222).

Esta perspectiva de ver e estudar a Arquivística, ao nosso entendimento, apesar de uma nova redação e de haverem quatro princípios, em nada fere o Princípio da Proveniência da forma que entendeu Michel Duchein e que expusemos no presente texto. Quando praticamos isomorfismos com outros saberes científicos, tal como o podemos fazer com a Lógica, poderemos ver que nas axiomáticas da Lógica Crisp há diferentes sistemas, ou formatos, em uma mesma teoria. Dopp (1970), em seu texto *Noções de Lógica Formal*, aponta diversas axiomáticas equivalentes em Lógica⁶²; umas com doze axiomas, outras com dez, quatro e assim por diante. Entretanto, apesar da enunciação ser diferente, ou mesmo quantitativamente diferente, não significa mudar de teoria, apenas de eleger enunciados de partida (axiomas) que formulam a mesma teoria de diferentes formas. No formato X, dos axiomas deste podem-se deduzir todos os axiomas do formato Y e o contrário também será possível no caso de ser uma mesma teoria. Este aprendizado de que formatos teóricos distintos podem significar uma mesma teoria só é possível quando nos damos ao trabalho de

⁶² Dopp (1970) denomina que tais axiomáticas são “sistemas” de Lógica. Aqui optamos pela terminologia “formatos” utilizado tal como em Vieira (2004).

buscar a compreensão por isomorfismos e de entender que todos os enunciados de uma são também dedutíveis dos enunciados da outra⁶³.

Vencida esta escaramuça metodológica-isomórfica, aponta-nos Schmidt (2012) que a perspectiva sistêmica nasce de discussões acadêmicas preocupadas com a inserção de uma fundamentação teórico-metodológica, valorizando o “saber” ao “mero fazer” arquivístico. Isto é, tendo em vista que o “fazer” sempre foi muito mais valorizado em comparação com o “saber”, haveria um re-balanceamento entre o “fazer” e o “saber”; naturalmente que o empírico como elemento de testabilidade (tal como em Popper, 1993) não perderá seu primado; entretanto, não seria o “mero fazer” o mote da Arquivologia como ciência, e sim o cumprimento dos critérios metodológico-organizacionais a que estão sujeitas as demais ciências (construção de teorias, construção de modelos e enunciados de previsão – quiçá controle – e testabilidade dessas por cotejamento com o empírico). Para não perdermos o significado do que seria inserir realmente uma perspectiva sistêmica, vejamos o que nos ensina Bertalanffy (1977):

O que pode ser obscurecido nesses desenvolvimentos – por mais importantes que sejam – é o fato de que a teoria dos sistemas consiste numa ampla concepção que transcende de muito os problemas e exigências tecnológicas, é uma reorientação que se tornou necessária na ciência em geral e na gama de disciplinas que vão da física e da biologia às ciências sociais e do comportamento e à filosofia. É uma concepção operatória, com graus variáveis de sucesso e exatidão, em diversos terrenos, e anuncia uma nova compreensão do mundo, de considerável impacto (BERTLANFFY, 1977, p.7-8).

Assim, apesar de haver alguns equívocos naqueles que queiram ver somente facilidades em inserir na Arquivologia uma perspectiva sistêmica, devemos ressaltar e apontar o próprio entendimento do “pai da Teoria Geral dos Sistemas”, isto é, os “ônus investigativos” que dessa *inserção de perspectiva sistêmica* vão advir:

Conforme dissemos, há vários enfoques para tratar desses problemas. Usamos intencionalmente a expressão um tanto vaga “enfoques” porque estes são logicamente não homogêneos, representam diferentes modelos conceituais, técnicas matemáticas, pontos de vista gerais, etc., concordando porém na qualidade de serem “teorias dos sistemas”. Deixando de lado os enfoques da pesquisa aplicada dos sistemas, tais como os sistemas de engenharia a pesquisa operacional, a programação linear e não linear, etc., os enfoques mais importantes são os seguintes (para uma boa compilação veja-se DRISHEL, 1968.)

⁶³ Eventualmente dois formatos teóricos diferentes possuem alguns axiomas semelhantes, ainda que difiram parcialmente; mas todos os enunciados de um determinado formato são verdadeiros no outro formato. Diferentemente no que acontece com duas teorias diferentes. Na Geometria de Lobachevsky *não é verdade* que “a soma dos ângulos internos de um triângulo seja de 180° (cento e oitenta graus)”, tal como é provado como enunciado verdadeiro na Geometria Euclideana.

A teoria “clássica” dos sistemas aplica matemática clássica, isto é, o cálculo. [...]

Computação e simulação. Os conjuntos diferenciais simultâneos como meio de “modelar” ou definir um sistema são, quando lineares, [...].

Teoria dos compartimentos. [...] Compreende-se bem que as dificuldades matemáticas tornan-se proibitivas no caso de sistemas de três ou mais compartimentos. As transformações de Laplace, [...].

Teoria dos gráficos. Muitos problemas dos sistemas referem-se a propriedades estruturais ou topológicas dos sistemas, e não a relações quantitativas [...]. Matematicamente, esta teoria liga-se à álgebra das matrizes e em forma de modelos [...].

A teoria das redes. Por sua vez, liga-se às teorias dos conjuntos, dos gráficos, dos compartimentos, etc. e aplica-se a sistemas tais como as redes nervosas [...].

A cibernética. [...] Em biologia e em outras ciências fundamentais o modelo cibernético serve para descrever a estrutura forma de mecanismos reguladores, por exemplo, por meio de diagramas de blocos e de fluxogramas.

A teoria da informação, no sentido de Shannon e Weaver (1949), baseia-se no conceito de informação, definido por uma expressão isomórfica à da entropia negativa da termodinâmica.

A teoria dos autômatos (veja-se MINSKY, 1967), é a teoria dos autômatos abstratos, com entrada, saída, possivelmente ensaios e erros e aprendizagem. Um modelo geral é a máquina de Turing (1936).

A teoria dos jogos (VON NEUMANN; MORGENSTERN, 1947) é um enfoque diferente mas pode ser classificado entre as ciências dos sistemas porque diz respeito aos comportamento de jogadores supostamente “racionais” [...].

A teoria da decisão é uma teoria matemática que trata de escolhas entre alternativas.

A teoria da fila refere-se à otimização de arranjos em condições de aglomeração (BERTLANFFY, 1977, p.38-43).

Assim, alerta-se aos adeptos de nomes sugestivos e de marketing, ou de frases de impacto, de que a inserção de uma teoria ou de uma perspectiva sistêmica na Arquivologia significaria ônus; isto é, cotejar minimamente a Arquivologia com algumas dessas opções de teorias dos sistemas que nos foi apresentada anteriormente por Bertlanffy. Particularmente, sugere-se a observação de teorias ou formatos da Lógica Crisp, já em uma forma estruturada, tal como, por exemplo, em Vieira (2004)⁶⁴ ou em Nolt e Rohatyn (1991) com o fito de

⁶⁴ A Sociedade Brasileira de Computação (SBC), por exemplo, recomenda nos currículos da área a inserção de Lógica (disponível em: <<http://bcc.ime.usp.br/principal/documentos/sbc2005.pdf>>. Acesso em: 06 jul. 2014). Vieira, um dos autores de Lógica citados no texto, lecionou por anos em curso de graduação na área de Ciência da Informação e, portanto, parte da experiência aposta em seu texto de 2004 é oriundo da experiência obtida no magistério de cursos ligados à Ciência da Informação.

compreender melhor as possibilidades de analogias por isomorfismo. Também o vanguardista Instituto Santa Fé (EUA) aponta a necessidade de se observarem certos investimentos em aprendizagem:

Eu acredito que estamos entrando em um período de *transciência*, onde está se tornando necessário que o treinamento em áreas nas quais os princípios fundamentais da matemática, computação e lógica deveriam ser emancipados de uma única classe de estudos de casos historicamente contingentes. Por exemplo, a física estatística vai continuar a ser útil para a compreensão dos fenômenos sociais, como, tradicionalmente, tem sido no estudo das propriedades da matéria condensada (KRAKAUER, 2011, p.2, tradução nossa)⁶⁵.

Assim, entendendo a Arquivística como uma disciplina científica voltada a uma esquematização-organização pela perspectiva da Teoria Geral dos Sistemas é mister observar o que é feito na adaptação ou construção das várias áreas de conhecimento que adotaram a mesma premissa. Imaginar que podemos abdicar de cotejar-nos com a Lógica, a Computação e a Matemática, tal como nos aponta o texto de Krakauer anteriormente transcrito, é perdermo-nos em um oceano de equívocos e não observar o que já nos é posto por outros grandes pensadores. Finalmente, antes de passarmos a uma releitura de Michel Duchein sob esta perspectiva sistêmica, vejamos mais um dos vários vaticínios do pensador espanhol José Ortega Y Gasset, pensador que cuidou de refletir também sobre a prática do bibliotecário em sua *Misión del bibliotecario* (1967). No *Em torno a Galileu: esquemas das crises* vaticina o referido autor sobre ser a ciência uma construção que coteja o empírico com a “imaginação” mental:

Pois bem, eu tenho a convicção de que se avizinha um esplêndido florescimento das ciências históricas, devido a que os historiadores se resolverão a fazer *mutatis mutandis*, frete aos fatos históricos o mesmo que Galileu começou frente aos físicos. Convencer-se-ão de que a ciência, entenda-se toda a ciência de coisas, sejam estas corporais ou espirituais, é tanto obra de imaginação quanto de observação, que esta última não é possível sem aquela – em suma, que a ciência é construção (ORTEGA Y GASSET, 1989, p. 27, negrito nosso).

Assim, necessária uma construção mental – teoria – e a validação das previsões da teoria (tal como antes apontamos em Stiglitz, Walsh e também em Popper) com a realidade empírica seja na História, na Sociologia, na Física ou na Arquivologia.

⁶⁵ O texto original de Krakauer (2011, p.2): “I believe we are entering a period of transcience, where it is becoming necessary that training in areas with fundamental mathematical, computational, and logical principles should be emancipated from a single class of historically contingent case studies. For example, statistical physics will continue to be every bit as useful in understanding social phenomena as it traditionally has been in studying properties of condensed matter”.

4 RELEITURA DE MICHEL DUCHEIN A PARTIR DA PERSPECTIVA SISTÊMICA

A abordagem sistêmica como um modelo teórico possível para a Arquivística permite, além de uma dimensão orgânica dos objetos, enfatizar a vertente do saber como também uma força propulsora para as atividades prático-empíricas da área. Assim fundamentado, propomos nesta parte do texto rever alguns dos problemas pontuados no artigo de Michel Duchein.

Ao aproximar a teoria sistêmica da Arquivologia teremos, não só uma perspectiva de construção de uma base teórica mais próxima de outras áreas mais consolidadas do conhecimento científico (cristalizadas a partir da observação organizacional da Lógica, da Teoria Geral dos Sistemas, da moderna computação e de exigências metodológicas tal como as apostas por Popper), como também aglutinar as características dos documentos de arquivos (imparcialidade, autenticidade, naturalidade, inter-relacionados e unicidade), tendo em vista que ao registrar os fatos jurídico-administrativos a veracidade de tais documentos efetua-se principalmente pela manutenção destas características e que são elementos fundamentais do pensamento sistêmico.

Portanto, torna-se quase um ato natural buscar-se ter na Arquivologia uma base teórica de *perspectiva sistêmica*⁶⁶, deixando, ainda assim, o primado do Princípio da Proveniência; ou, *alternativamente*, buscando-se outro formato, tal como em Silva (et al., 1999). Reforçamos aqui que: ao nosso entendimento, não há nova teoria com a proposta de Silva (et al., 1999), dado que *este não exibiu* nenhum enunciado contraditório aos que se deduzem da teoria atual. Assim, em ambas as propostas (formatos) teóricas vemos: seja de forma direta (admissão por escolha do Princípio da Proveniência), ou seja de forma indireta (sendo o Princípio de Proveniência dedutível de novos princípios) verifica-se a permanência do Princípio da Proveniência como elemento da construção teórica em Arquivologia.

Assim, em um ou outro formato teórico, conseguimos antever para a Arquivologia um corpo científico (teoria, modelos de previsão, testabilidades, etc.) para além dos aspectos meramente técnico-operacionais desta área do conhecimento.

Ao utilizarmos uma perspectiva sistêmica podemos observar mais facilmente as interligações entre as partes que formam o todo, analisando o sistema informacional tanto de um ponto de vista macro como de um ponto de vista micro. Quanto a escolher entre as

⁶⁶ *Perspectiva Sistêmica* como possibilidade de organização de um saber científico com base na observação das várias subteorias já consolidadas em Teoria Geral dos Sistemas (verificação de possibilidade de isomorfismos), tal como as subteorias apontadas em transcrição anterior do presente texto (BERTLANFFY, 1977, p.38-43).

posições estanques maximalista ou minimalista, escolhas que no texto de Duchein, não passavam por uma análise teórica, aqui tal escolha deixa de sê-lo e passa a ser decisão tomada com base nas premissas que o profissional de arquivo partirá para deduzir a melhor escolha estas duas posições. Note-se que, além dos Princípios da Arquivística, existem “regras ou situações contingenciais”. Os ordenamentos jurídicos do Brasil e do Canadá, por exemplo, são diferentes, e seria ingênuo pensar que tudo o que é normatizado em um dos países, sobre documentação, também o será no outro (e ainda a as normatizações discricionárias dos órgãos e gestores públicos). Não obstante as questões meramente legais, há ainda questões que deverão estar no polo das premissas: custos, dotação orçamentária, condições das edificações disponibilizadas para arquivo (pode ocorrer que a opção maximalista possua melhores condições de guarda e manutenção dos documentos arquivísticos que a perspectiva minimalista que poderia colocar o a preservação dos documentos arquivísticos em risco de não ser minimamente satisfatória), etc.. Assim, a subjunção (previsão tal qual vimos no exemplo de Stiglitz e Walsh) que acarreta a assunção da posição maximalista ou minimalista no caso em concreto (decisão que estará no conseqüente da subjunção) deverá ser deduzida do corpo teórico e das premissas contingentes. Pesos e contrapesos em situações que vão integrar o conjunto de premissas (condições, cláusulas, etc.) são bastante comuns justamente na área de Engenharia de Sistemas. Notório que para haver decisões técnico-científicas por parte do profissional de Arquivologia não necessariamente as inferências terão caráter quantitativo, pois se infere (deduz-se, segundo o Método Hipotético-Dedutivo) em qualquer ciência humana ou social.

Assim, satisfeitas algumas pequenas questões de discurso e alguns necessários balanceamentos entre o empírico e o teórico, tal qual já ocorreu em muitas outras áreas do conhecimento científico, então não avistamos nesse momento qualquer discrepância entre a manutenção do Princípio da Proveniência e a simultânea adoção de uma perspectiva sistêmica para a construção e desenvolvimento da Ciência Arquivística. Note-se que na *adoção de uma perspectiva sistêmica* em Arquivologia devemos observar o que já ocorreu em certos outros campos. Vejamos o que nos aponta Bertalanffy (1977):

De fato, conceitos, modelos e leis similares apareceram muitas vezes em campos bastante diferentes, independentemente e com base em fatos totalmente diferentes. Há muitos casos em que princípios idênticos foram descobertos várias vezes porque os pesquisadores que trabalhavam em um campo ignoravam que a estrutura exigida já estava desenvolvida em outro campo (BERTALANFFY, 1977, p.56).

A adaptação de teorias para uma perspectiva sistêmica já ocorreu em outros campos do saber, está na Biologia, na Matemática, na Química, em vários domínios da Física, nas

engenharias, na Ciência da Computação, na Economia, etc. Por outro lado, devemos alertar que não serão analogias vagas, entendidas aqui uma analogia sem haver cotejamento com outras construções teóricas, ou modelos, já consolidados que nos serão úteis e, nesse sentido, o próprio Bertalanffy (1977) nos alerta:

Parece, portanto, que uma teoria geral dos sistemas seria um instrumento útil capaz de fornecer modelos a serem usados em diferentes campos e transferidos de uns para outros, salvaguardando ao mesmo tempo do perigo das analogias vagas, que muitas vezes prejudicaram os progressos nesses campos (BERTALANFFY, 1977, p.57).

Diante do exposto, infere-se que havendo adoção de práticas interdisciplinares, da inclusão da perspectiva sistêmica e da observação acurada de outras estruturas isomórficas, então haverá uma aceleração na direção da Ciência Arquivística.

5 CONSIDERAÇÕES FINAIS

A introdução da perspectiva sistêmica em Arquivologia, em regra, não nos obriga necessariamente a assumir qualquer enunciado científico que torne o *novo sistema teórico da Arquivística, sob a perspectiva sistêmica, inconsistente*. Assim, plenamente possível à permanência do Princípio da Proveniência junto à forma sistêmica. A perspectiva sistêmica é um conjunto de modos (tal como nos alertou Bertalanffy sobre a variedade de possibilidades de teorias como: teoria clássica dos sistemas, computação e simulação, teoria dos compartimentos, teoria dos gráficos, a teoria das redes, etc.) de organização de conhecimentos onde seu “objeto é a formulação de princípios válidos para os ‘sistemas’ em geral, qualquer que seja a natureza dos elementos que os compõem e as relações ou forças existentes entre eles” (BERTALANFFY, 1977, p.61). Ainda nesse sentido, o próprio Bertalanffy nos corrobora que (BERTALANFFY, 1977, p. 61) “em forma elaborada seria uma disciplina lógico-matemática, em si mesma puramente formal, mas aplicável às várias ciências empíricas”. Desta forma, entendemos que o cotejamento com o modelo lógico do Método Hipotético-Dedutivo⁶⁷ seja o caminho natural para tal entronização de uma perspectiva sistêmica em Arquivologia⁶⁸.

Fundamental é pontuar que nesses desenvolvimentos que virão, tal como estamos aqui antevendo, devemos reconhecer a enorme contribuição de nomes como Silva (1999), Belotto (2006), Duchein (1986), Albuquerque (2013), Rosseau; Couture (1998), Rodrigues (2011),

⁶⁷ Tal como exibido em Vieira (2004).

⁶⁸ Vários cursos de Biblioteconomia já possuem Lógica como disciplina obrigatória, pecando apenas em manter departamentos de Filosofia ministrando-as sem, portanto, haver cotejamento com os formatos mais modernos da Lógica Crisp e reportando-se a até mesmo a “formas Aristotélicas” para seu ensino.

Heredia Herrera (1993), Cook (2001), Martín-Pozuelo Campillos (1996), etc. (aos quais se somam vários outros não nominados aqui, e por tal omissão nos desculpamos), que vem desenvolvendo a Arquivologia tanto como ciência quanto como técnica⁶⁹ e que devemos render as devidas homenagens pelas valiosas contribuições ao campo.

Assim sendo, conclui-se que a construção de uma *teoria tipo sistêmica*⁷⁰ para a Arquivologia permitirá uma melhor organização e construção do conhecimento arquivístico enquanto ciência e técnica. Nesta organização e construção, se admitirmos como premissa a realização de isomorfismos com outros campos do saber científico (tal como é usual na perspectiva sistêmica) será, conseqüentemente, baseada na eleição de princípios e axiomas, na construção de conceitos e no desenvolvimento de outros mecanismos teórico-metodológicos⁷¹ que permitam a referida organização dos documentos arquivísticos para a recuperação, armazenamento e disponibilização da informação e, ao menos por enquanto (dado que não se verificou, em regra, qualquer contradição aparente com a adoção do sistemismo tal como exposta por Bertalanffy), possuirá ainda em seu primado o Princípio da Proveniência como princípio norteador e basilar desta ciência em formação.

A estruturação de uma disciplina como área do conhecimento científico e como profissão é bem mais factível na medida em que estejam consolidadas teoria, metodologia, prática, terminologia, legislação e formação profissional. Portanto, o caminho a percorrer, ao nosso entendimento, é investirmos como comunidade acadêmica de Ciência da Informação e, em particular, como pesquisadores da Arquivologia, no aperfeiçoamento de nossa visão sistêmica e na prática intensa de realizarmos análises isomórficas com outros campos do saber. Tal investimento interdisciplinar e, ousamos dizer, com vistas à transciência, é algo que outros pesquisadores já avistaram e que percebemos ser o caminho natural desta jovem Ciência da Informação.

REFERÊNCIAS

ACEVES, Merizanda, Ramírez. La Archivística: ¿ciencia o técnica? In: RENDÓN ROJAS, Miguel Ángel (Coord.). **Bibliotecología, Archivística, Documentación: intradisciplina, interdisciplina o transdisciplinaridad**. Mexico: UNAM, Centro Universitario de Investigaciones Bibliotecológicas, 2011.

⁶⁹ Como também acontece em Medicina, Direito, nas várias engenharias, inclusive na Engenharia de Produção (de Métodos) e na Engenharia de Sistemas, entre outros campos do saber científico, mas que também, ao mesmo tempo, são técnicas.

⁷⁰ Pois pode ocorrer, com a introdução do sistemismo, o surgimento de mais de uma teoria como ocorre em outros campos do conhecimento científico.

⁷¹ Nesses mecanismos incluem-se critérios próprios (do campo específico de Arquivologia) associados à validação ou refutação nos procedimentos de testabilidade.

ALBUQUERQUE, Ana Cristina de; SOUTO, Diana Vilas Boas. Acerca do princípio da proveniência: apontamentos conceituais. In: **Ágora**, v. 23, p.14-44, 2013. Disponível em: <http://agora.emnuvens.com.br/ra/article/view/425/pdf_8>. Acesso em: 6 jun. 2014.

BELLOTTO, Heloísa Liberalli. **Arquivos permanentes: tratamento documental**. Rio de Janeiro: FGV, 2006.

BERTALANFFY, Ludwig von. **Teoria Geral dos Sistemas**. Rio de Janeiro: FGV, 1977.

COOK, Terry. Archival Science and Postmodernism: new formulation for old concepts. In: **Archival Science**, v. 1, n. 1, p.3-24, 2001.

DOPP, Joseph. **Noções de lógica formal**. São Paulo: Herder, 1970.

DUCHEIN, Michel. O Respeito aos Fundos em Arquivística: Princípios teóricos e problemas práticos. **Arquivo & Administração**, Rio de Janeiro, v. 10-14, n. 1, abr. 1982 e ago. 1986.

HEREDIA HERRERA, Antonia. **Archivística general: teoria y práctica**. Sevilla: Servicio de Publicaciones de la Diputación de Sevilla, 1993.

KRAKAUER, David C. Transcience: disciplines and the advance of plenary knowledge. In: **SFI Bulletin**, Santa Fé Institute, Santa Fé (EUA), 2011. Disponível em: <http://www.santafe.edu/media/bulletin_pdf/sfi_bulletin_spring_2011.pdf>. Acesso em: 6 jun. 2014.

KUHN, T. S. **A estrutura das revoluções científicas**. São Paulo, Ed. Perspectiva, 1990.

MARTÍN-POZUELO CAMPILLOS, Maria Paz. **La construcción teórica em Archivística: el principio de procedência**. Madrid: Universidad Carlo III. 1996.

NOLT, John; ROHATYN, Dennis. **Lógica**. São Paulo: MacGraw-Hill, 1991. (Coleção *Schaum*).

ORTEGA Y GASSET, José. **Misión del bibliotecario**. Madrid: Revista de Occidente, 1967.

POPPER, Karl R. **A Lógica da pesquisa científica**. São Paulo, Ed. Cultrix, 1993.

POPPER, Karl R. **Conjecturas e refutações**. Brasília: UnB, 1972.

QUINE, Willard Van Orman. **Filosofia da Lógica**. Rio de Janeiro: Zahar, 1972.

RODRIGUES, Ana Célia. Identificação arquivística: subsídios para a construção teórica da metodologia na perspectiva da tradição brasileira. In: **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 04, p.44-81, 2011. Disponível em: <<http://inseer.ibict.br/ancib/index.php/tpbci/article/view/44/81>>. Acesso em: 20 mar. 2014.

ROUSSEAU, Jean-Yves; COUTURE, Carol. **Os Fundamentos da Disciplina Arquivística**. Lisboa: Publicações Dom Quixote, 1998.

SCHMIDT, Clarissa Moreira dos Santos. **Arquivologia e a construção do seu objeto científico: concepções, trajetórias, contextualizações**. São Paulo: Escola de Comunicações e Artes (ECA/USP). Tese (Doutorado em Comunicação) ECA/USP, 2012.

SILVA, Armando Malheiro da; RIBEIRO, Fernanda; RAMOS, Júlio; REAL, Manuel Luís. **Arquivística Teoria e prática de uma Ciência da Informação**. Porto: Edições Afrontamento, 1999.

SOCIEDADE BRASILEIRA DE COMPUTAÇÃO. **Currículo de referência da SBC para cursos de graduação em Bacharelado em Ciência da Computação e Engenharia de Computação**. SBC, São Paulo, 2005. Disponível em: <<http://bcc.ime.usp.br/principal/documentos/sbc2005.pdf>>. Acesso em: 6 jul. 2014.

SOUSA, Renato Tarciso Barbosa de Sousa. A Classificação como função matricial do que-fazer arquivístico. In: SANTOS, Vanderlei Batista dos. **Arquivística temas contemporâneos: classificação, preservação digital, gestão do conhecimento**. Brasília: SENAC, 2007. p. 79-172.

STIGLITZ, Joseph Eugene; WALSH, Carl Eugene. **Introdução à Macroeconomia**. Rio de Janeiro: Campus, 2003.

VIEIRA, A. A. N. **Lógica: método semiótico estruturado**. Rio de Janeiro, Sarau Cultural, 2004.

OS REGISTROS DE AUTORIDADE EM SISTEMAS DE INFORMAÇÃO: UMA PERSPECTIVA BIBLIOTECÔNOMICA E ARQUIVÍSTICA

AUTHORITY RECORDS IN INFORMATION SYSTEMS: ONE LIBRARIANSHIP AND ARCHIVAL PERSPECTIVE

Dunia Llanes Padrón
Mariângela Spotti Lopes Fujita
Flávia Maria Bastos

Resumo: Nesta pesquisa se apresenta uma perspectiva geral da criação de registros de autoridade em biblioteca e arquivos. Para isso, apresenta-se, em primeiro lugar, uma reflexão sobre os registros de autoridade nesses sistemas de informação, para, em segundo lugar, expor uma comparação entre ambos os registros, partindo das semelhanças e diferenças existentes nessas entidades e a realização desta operação. Finalmente são abordadas as considerações finais sobre o tema de pesquisa.

Palavras-chave: Registros de autoridade em sistemas de informação. Formas autorizadas de nomes. Controle de autoridades. Tipos de registros de autoridade.

Abstract: This research is presented as an overview of the creation of authority registers in libraries and archives. For this, it is shown first, a reflection on authority records in these information systems. Secondly, it presents a comparison between the two records based on the similarities and differences in these institutions in carrying out this operation. Finally, the final considerations on the investigation subject are approached.

Keywords: Authority records information systems. Authorized name forms. Control of authorities. Types of authority record.

1 INTRODUÇÃO

As bibliotecas e os arquivos sempre tiveram a obrigação de facilitar às pessoas a consulta de documentos e o acesso a eles. Em vistas disso, a representação da informação tem sido uma das tarefas prioritárias para os profissionais da informação. A função da descrição, em instituições de informação, tem sido promover e proporcionar o uso dos documentos.

O desenvolvimento das tecnologias e as comunicações obrigaram a transformar o modo como os profissionais da informação devem apresentar e facilitar esse acesso. Os processos de representar e de recuperar a informação mudaram, pois os atuais usuários modificaram sua maneira de pesquisa. Os novos usuários desejam contar com ferramentas de recuperação mais homogêneas e melhor estruturadas e obter, em uma só pesquisa, todas as fontes de informação sem estar fisicamente em uma instituição e sem ter que navegar por muitas páginas da internet.

É nesse contexto que se deve entender a criação e a manutenção, na atualidade, dos registros de autoridade nos sistemas de informação, em que o impacto tecnológico também tem variado as concepções. As autoridades constituem um dos pontos de acesso mais importantes na recuperação da informação. O usuário procura, identifica e acessa a

informação por meio de nomes de pessoas, instituições, famílias, obras, matérias ou lugares geográficos. Neste sentido, o controle de autoridades e a normatização de cabeçalhos ou formas autorizadas de nomes constituem aspectos chave para a recuperação, o intercâmbio, a difusão e o uso da informação.

Nos últimos anos, a comunidade internacional de bibliotecários e arquivistas têm trabalhado, arduamente, para fortalecer as concepções teóricas desse tema e, assim, tem desenvolvido normas e diretrizes para sistematizar a representação de autoridades e a criação dos registros. Tudo isso facilita o trabalho dos profissionais e permite aos usuários identificar as fontes de informação que cobrem suas necessidades. Além disso, proporciona a integração de bibliotecas e de arquivos nas novas dinâmicas da sociedade, caracterizada pela presença incondicional das tecnologias.

Essas novas realidades nos levaram a estudar a criação de registros de autoridade em bibliotecas e arquivos; Conseqüentemente, esse assunto tornou-se o principal objeto de pesquisa. Portanto, o problema gira em torno da seguinte questão:

Quais são as semelhanças e as diferenças que existem entre os registros de autoridade das bibliotecas e os registros de autoridade dos arquivos?

Para ajudar a delimitar o alcance do trabalho, respondendo à pergunta enunciada, foram determinados os seguintes objetivos:

Estudar os principais aspectos teóricos e normativos relacionados com a criação de registros de autoridade em bibliotecas e arquivos.

Discutir as semelhanças e diferenças entre ambos os registros de autoridade.

A pesquisa apresentada pretende saber quais são os novos paradigmas e as práticas na constituição de sistemas de autoridades e, por conseguinte, o intercâmbio de dados de autoridades entre instituições de informação internacionais e nacionais.

2 OS REGISTROS DE AUTORIDADE EM BIBLIOTECAS

No contexto das bibliotecas, a catalogação é considerada um processo comunicativo que envolve a descrição de um item e a escolha dos pontos de acesso a um registro bibliográfico durante a recuperação de informações.

As autoras Mey e Silveira (2009, p.96) compartilham que um registro bibliográfico é “o conjunto completo da representação, com todas as suas partes [descrição bibliográfica, pontos de acesso e dados de localização]”.

Para Campos (2003, p.2), o registro bibliográfico é representado pelo ponto de acesso que direciona o usuário para o registro; a descrição bibliográfica possibilita ao usuário decidir

se o recurso interessa ou não; e a localização encaminha o usuário para o documento desejado.

Conforme os Princípios Internacionais de Catalogação (2009), os pontos de acesso, como parte dos registros bibliográficos, delimitam os resultados da busca, possibilitando a recuperação confiável dos registros bibliográficos bem como de registros de autoridades e seus respectivos recursos bibliográficos associados.

O controle dos pontos de acesso é de fundamental importância na recuperação dos registros bibliográficos. Essa atividade também associa as expressões de controle de autoridade e trabalho de autoridade. A necessidade de se controlarem as formas dos nomes bem como as metodologias para a produção de registros uniformes foi discutida na *International Federation of Library Associations and Institutions - IFLA* (2009), que considerou os pontos de acesso como elementos tanto do registro bibliográfico quanto do registro de autoridade.

As diversas iniciativas e a publicação de várias diretrizes gerais para orientar a criação e a representação dos pontos de acesso também são refletidas no trabalho de normalização de construção dos pontos de acesso, o que deu origem à criação de um Grupo de Estudo sobre o Sistema de Autoridade Internacional com dois objetivos declarados: definir os elementos de um registro de autoridade e preparar um formato de intercâmbio de dados. Os resultados do trabalho desse grupo foram definidos em 1984, com a publicação das *Guidelines for Authority and Reference Entries – GARE*.

Em 2001, foi editada uma segunda edição, intitulada *Guidelines for Authority Records and References – GARR*, com escopo para especificar requisitos para a exibição de informações relativas a cabeçalhos autorizados e a referências em listas de autoridades, sejam em materiais impressos ou eletrônicos. O objetivo do GARR foi preencher a lacuna existente em códigos e normas de catalogação para a construção de registros de autoridade e, para isso, apresenta sete áreas de dados para a construção do registro de autoridade (DIRECTRICES, 2001, p.42⁷² *apud* OLIVEIRA, 2013, p.37):

Área 1 – Entrada autorizada

- Contém a entrada autorizada estabelecida pela agência catalogadora responsável. A entrada pode ser uma pessoa, entidade ou obra. Pode incluir também entradas paralelas.

Exemplos:

⁷² DIRECTRICES para registros de autoridad y referencias. 2 ed. rev. por el Grupo de Trabajo de la IFLA para la Revisión de las GARE. IFLA, 2001. Disponível em: <<http://www.ifla.org/files/assets/cataloguing/garr/garr-es.pdf>>. Acesso em: 15 jul. 2014.

Nome pessoal: Assis, Machado de, 1839-1908

Nome entidade: Universidade de Brasília

Obra: Brasil. Código Civil <2002>

Área 2 – Nota de informação

- Nota de informação é inserida quando for necessário explicar a relação da entrada autorizada e as entradas remissivas ou relacionadas. Trata-se de notas mostradas ao público. Podem conter uma explicação sobre as relações entre duas ou mais entradas de nomes pessoais; breve história de uma organização falando sobre mudanças de nome, etc.; uma explicação da relação entre duas ou mais obras, ou entre partes da obra e a sua entrada geral, ou informações que têm como objetivo identificar a entidade que dirige.

Área 3 – Referência “Ver”

- A referência “Ver” mostra as relações entre as formas variantes da entrada autorizada e serve para guiar o usuário até a forma autorizada.

- Exemplo:
Assis, Machado de, (1839-1908)
<Assis, Joaquim Maria Machado de

Área 4 – Referência “Ver também”

- A referência “Ver também” mostra a relação da entrada autorizada com outras também autorizadas levando o usuário a consultar também essa outra entrada.

- Exemplo:
Tahan, Malba, (1895-1974)
<< Souza, Julio Cesar de Mello e, (1895-1974)

Área 5 – Nota do catalogador

- Essas notas não são normalmente exibidas ao público e servem como ajuda a quem cataloga. Referem-se a como a entrada autorizada deve ser usada, de onde as informações foram extraídas, quais regras foram aplicadas, etc..

Área 6 – Fonte

- Identifica a agência catalogadora responsável pela entrada, as regras de catalogação que foram usadas e as datas.

Área 7 – ISADN

- Identifica a numeração ISADN, que, como foi dito anteriormente, teve sua implementação abandonada.

As diretrizes apresentadas nas GARR também se refletiram na composição dos formatos de autoridade, como o *United International Machinal Readable Cataloging* (UNIMARC). No caso do formato *Machinal Readable Cataloging 21* (MARC 21), constata-se que sua estrutura é semelhante à proposta das GARR no que se refere à apresentação dos registros. Dessa forma, permite que áreas das GARR tenham correspondência nos campos estabelecidos pelo padrão MARC 21.

Durante o processo de recuperação de informações, o usuário passa a ser guiado a materiais existentes em catálogos de bibliotecas por meio do registro de autoridade, o que proporciona acesso uniforme a partir do estabelecimento das formas de nomes - para pessoas (TAB. 1), organizações (TAB. 2), lugares (TAB. 3), títulos (TAB. 4) e assuntos (TAB. 5 e 6) usados em registros bibliográficos como se demonstra nas tabelas abaixo:

TABELA 1 - Nome pessoal

<i>CAMPO</i>	<i>DADO</i>
FMT	AU
LDR	nz 2200157n 4500
001	UNAU000107035
003	BBD
008	951127nn acnnaabn a aaa d
040	la UNESP lb por
1001	la Freyre, Gilberto ld 1900-1987
4001	la Freire, Gilberto ld 1900-1987
4001	la Freire, Gilberto de Melo ld 1900-1987
4001	la Freyre, Gilberto de Melo ld 1900-1987
665	la Casa grande & senzala, 1978
670	la Koogan, 1982 (Freyre, Gilberto de Melo) lb DN
670	la Dic. Hist.-Biog. Bras., 1984 (Freire, Gilberto) lb DN
670	la J. Brasil, 19/07/87 lb DN
	000107035

Fonte: Catálogo Bibliográfico Online Athena da UNESP

O registro de autoridade é parte do controle de autoridade, uma das fases da catalogação descritiva e catalogação de assunto, que compreende o trabalho de autoridade determinado por um conjunto de operações guiadas pelas regras das *Anglo-American Cataloguing Rules, second edition* (AACR2), as quais estabelecem as entradas autorizadas e as remissivas de nome sem, contudo, estabelecer regras para construção de um registro de autoridade.

TABELA 2 - Nome corporativo

<i>CAMPO</i>	<i>DADO</i>
<i>FMT</i>	<i>AU</i>
<i>LDR</i>	<i>nz 2200157n 4500</i>
<i>001</i>	<i>UNAU000119513</i>
<i>003</i>	<i>BR-SpUEP</i>
<i>005</i>	<i>20140406220046.0</i>
<i>008</i>	<i>880916nn acznaabn a ana d</i>
<i>035</i>	<i>la (BR-RjFGV)CA000588241</i>
<i>040</i>	<i>la UNESP/BGI lb por</i>
<i>1102</i>	<i>la Universidade Estadual Paulista "Júlio de Mesquita Filho"</i>
<i>4102</i>	<i>la Universidade Estadual Paulista Júlio de Mesquita Filho</i>
<i>4102</i>	<i>la UNESP</i>
<i>4102</i>	<i>la Universidade Estadual Paulista</i>
<i>4102</i>	<i>la Universidade Júlio de Mesquita Filho</i>
<i>4102</i>	<i>la UNESP - Univ Estadual Paulista</i>
<i>665</i>	<i>la Criada pela Lei n.952, de 30/01/1976 como entidade autárquica</i>
<i>670</i>	<i>la Cab. unif. ent. col., 1979 lb V</i>
<i>670</i>	<i>la Cat. geral inst. ens. sup., 1986 lb V</i>
<i>670</i>	<i>la Perfil adm. publ. paulista, 1977 lb V</i>

670	la Res. Unesp 72/2009 lb (rem., Estabelece padrão para citação institucional da Universidade Estadual Paulista “Júlio de Mesquita Filho” em todas as publicações científicas nacionais e estrangeiras)
COR	la Universidade Estadual Paulista Júlio de Mesquita Filho \0 ZZZ
SYS	000119513

Fonte: Catálogo Bibliográfico Online Athena da UNESP

TABELA 3 - Área geográfica

CAMPO	DADO
FMT	AU
LDR	nz 2200085n 4500
001	UNAU000209263
003	BBD
008	900108 e adannbabn a ana d
040	la UNESP lb por
043	la s-bl-- lb s-bl-sp \2 BIRjFGVB
151	la São Paulo (Estado)
SYS	000209263

Fonte: Catálogo Bibliográfico Online Athena da UNESP

TABELA 4 - Título

CAMPO	DADO
*****	Documento 1 de 1
FMT	AU
LDR	nz a n 4500
001	UNAU000279559
003	BR-SpUEP
005	20140801121901.0
008	140801nn acznnaabn a ana d
035	la (BR-RjBN)2002020416213379med
040	la UNESP/BGI lb por
130 0	la Nitheroy, revista brasiliense
430 0	la Niterói, revista brasiliense
430 0	la Niterói, revista brasiliense de ciências, letras e artes
670	la Assunto de: Pinassi, Maria Orlanda. Três devotos, uma fé, nenhum milagre. 1998 lb (p. 11: Revista publicada em Paris, em 1836, por Gonçalves de Magalhães, Araújo Porto Alegre e Torres Homem; n.c.: Nitheroy, revista brasiliense de ciências, letras e artes)
670	la LC on line obras lb (Niterói, revista brasiliense)
670	la Delta, v.8 lb (p.[4813]: Niterói, revista de que saíram apenas dois números, editada em Paris, França)
OWN	la AUTREV
CAT	la ROSANE lb 40 lc 20140801 ll UEP10 lh 1219
SYS	000279559

Fonte: Catálogo Bibliográfico Online Athena da UNESP

TABELA 5 - Termo tópico

CAMPO	DADO
FMT	AU
LDR	nz 2200193n 4500

001	UNAU000217994
003	BBD
008	960628 e adannbabn a ana d
040	la UNESP lb por
150	la Ciência da informação
550	la Banco de dados bibliográficos
550	la Comunicação
550	la Documentação
550	la Processamento eletrônico de dados
550	la Recuperação da informação
550	la Serviços de informação
550	la Sistemas de recuperação da informação
667	la Pode ser subdividido geograficamente
667	la Para o Brasil usar a forma direta
750 0	la Information science l5 DLC
SYS	000217994

Fonte: Catálogo Bibliográfico Online Athena da UNESP

TABELA 6 - Termo tópico e subcabeçalho

CAMPO	DADO
FMT	AU
LDR	nz 2200145n 4500
001	UNAU000223734
003	BBD
008	920214 e adannbabn a ana d
040	la UNESP lb por
150	la Legislação
360	li para assuntos individuais usar Legislação como subcabeçalho conforme os exemplos la Fábricas -- Legislação; Florestas -- Legislação
667	la Pode ser subdividido geograficamente
667	la Para o Brasil usar a forma direta
680	li Aqui entram os trabalhos sobre o processo legislativo
750 0	la Law and legislation l5 DLC
750 0	la Legislation l5 DLC
SYS	000223734

Fonte: Catálogo Bibliográfico Online Athena da UNESP

A esse respeito, ressalta-se que os registros de autoridade da *Library of Congress Subject Headings* (LCSH) são constituídos com base nas regras AACR2 de construção do cabeçalho e também seguem as *Guidelines for Authority Records and References* (GARR) da IFLA para construção do registro.

O trabalho de autoridade de acordo com Assumpção e Santos (2011), envolve cinco elementos: pontos de acesso, dados de autoridade, registros de autoridade, arquivos de autoridade e sistemas de autoridade.

O controle de autoridade existe para determinar uma forma autorizada para cada entrada e estabelecer também as formas não autorizadas. Dessa forma, o processo consiste no

estabelecimento de uma só forma autorizada para os pontos de acesso em um catálogo e na explicitação das relações que vinculam os pontos de acesso entre si. Outra característica presente no controle de autoridade é sua consistência e, por isso, as definições de controle de autoridade giram em torno da normalização dos pontos de acesso, mas também tratam das relações existentes entre eles.

A importância do controle de autoridade e de sua aplicação ao catálogo está em, fornecer ferramentas para o usuário recuperar informação com precisão, reunir registros por suas similaridades e navegar satisfatoriamente pelos resultados obtidos (BOERIS, 2011). O controle de autoridade, portanto, é fundamental para a catalogação descritiva e a catalogação de assuntos, pois assegura a precisão na recuperação da informação.

O registro de autoridade, no contexto das bibliotecas, mantém relação com controle de autoridade. Campos (2003) corrobora essa idéia, ao afirmar que a formulação dos pontos de acesso deve ser normalizada, inserindo aí o controle de autoridade. Além disso, assegurar ao usuário a localização de um registro e sua capacitação para agrupar registros com características em comum, é necessária a normalização na construção dos pontos de acesso, ou seja, o controle de autoridade.

Em decorrência, o resultado do processo do controle de autoridade será o registro de autoridade, entendido como o conjunto de informações padronizadas sobre um ponto de acesso. O registro de autoridade é composto ao se estabelecer a forma do nome a ser utilizada, que é a autorizada, e suas referências imprescindíveis. Em um registro de autoridade de assuntos, por exemplo, é possível observar informações sobre a determinação da entrada autorizada, notas e o uso de cabeçalho não autorizado.

TABELA 7 - Registro de Autoridade de Assunto

<i>CAMPO</i>	<i>DADO</i>
*****	Documento 1 de 1
Formato	AU
Líder	nz 2200313n 4500
No. de controle	UNAU000251263
ID no. de controle	BBD
Última atualização	20130520150757.0
Dados fixos	901112 n anznnbabn a ana d
Fonte catalogadora	UNESP por
Cabeçalho-Termo to	Política internacional [entrada autorizada]
Ref. Compl. VER	Também as subdivisões - Relações exteriores e Política e governo sob
TAM	nomes de países [Nota]
VER-Termo tópico	Política mundial [termo não autorizado]
VER-Termo tópico	International politics [termo não autorizado]
Ver tbém - Termo t	h Cristianismo e problemas internacionais

Ver tbém - Termo t	h Islamismo e política internacional
Ver tbém - Termo t	h Política afro-asiática
Ver tbém - Termo t	h Potências médias
Ver tbém - Termo t	Geografia política
Ver tbém - Termo t	História política
Ver tbém - Termo t	Geopolítica
Ver tbém - Termo t	Organização internacional
Ver tbém - Termo t	Propaganda internacional
Ver tbém - Termo t	Questão oriental
Ver tbém - Termo t	Reconhecimento de Estado (Relações internacionais)
Ver tbém - Termo t	Relações internacionais
Ver tbém - Termo t	g Ciência política
Ver tbém - Termo t	g História universal
Nota geral pública	Obras que tratam da história política em geral e as razões históricas das relações entre as nações entram em História política. Obras sobre a teoria, entram em Relações internacionais. Obras sobre relações exteriores de países específicos, entram sob o nome do país com a subdivisão - Relações exteriores
Link-termo tópico	World politics
Owner	AUTREV
SYS	000251263

Fonte: Catálogo Bibliográfico Online Athena da UNESP

De acordo com (CAMPOS, 2003, p.3), a função do registro de autoridade abrange seis aspectos:

- Registra a forma normalizada de cada ponto de acesso.
- Possibilita o agrupamento de todos os registros com o mesmo ponto de acesso.
- Garante a normalização nos registros no catálogo.
- Documenta as decisões tomadas e as fontes utilizadas para estabelecer os pontos de acesso.
- Registra todas as formas do ponto de acesso e não apenas a escolhida como normativa.
- Documenta outras formas usadas para o ponto de acesso (como formas anteriores do nome) de modo a servir de orientação para o catalogador.

Esses aspectos asseguram uniformidade e sistematização ao trabalho do catalogador na construção e na manutenção do catálogo de registros bibliográfico com a finalidade de aprimoramento na recuperação da informação solicitada pelo usuário em sua estratégia de busca. Mais importante ainda é que a existência dos registros de autoridades garante a contínua correção automática de qualquer registro bibliográfico no que se refere à descrição e aos assuntos sem que seja necessário o catalogador se preocupar com o controle, refazer seu trabalho ou realizar inúmeras consultas sobre a correção dos pontos de acesso a cada vez que se insere um registro bibliográfico na base de dados. Construir os registros de autoridade é

fundamental no trabalho de normalização do catálogo, tanto para o catalogador quanto para o usuário.

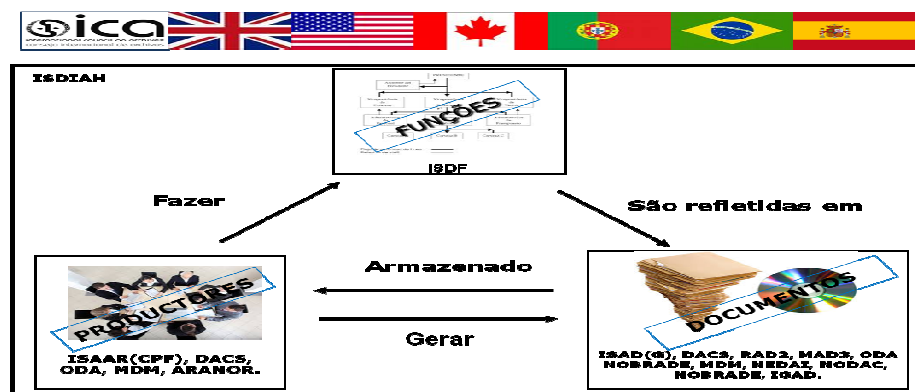
3 OS REGISTROS DE AUTORIDADE NOS ARQUIVOS

O registro de autoridade em arquivo é um conceito que deve ser entendido e abordado desde o princípio da descrição arquivística e sua normatização. A descrição arquivística é um dos processos intelectuais mais importantes que se realiza nos arquivos, visto ser a ponte que une os usuários com os acervos documentais e seu propósito é “identificar e explicar o contexto e o conteúdo do material arquivístico a fim de promover sua acessibilidade” (ISAD(G), 2000, p.12).

No contexto atual da arquivologia, esse processo tem sofrido variações nas estratégias de trabalho, vinculadas, principalmente, ao processo internacional de normatização. As mudanças produzidas na sociedade de informação, marcada pelas novas tecnologias, a globalização e a diversidade de usuários levaram os órgãos arquivísticos internacionais e as comunidades de cada país a desenvolver normas de estrutura de dados, de conteúdo, de apresentação e de formato que se adaptaram aos requerimentos dos usuários e às práticas arquivísticas.

Como resultado têm-se desenvolvido novas normas internacionais, nacionais e regionais que, em conjunto, são utilizadas para descrever documentos, produtores, funções e instituições arquivísticas. Na FIG. 1, mostra-se, gráfica e resumidamente, quais são as normas existentes para descrever as entidades arquivísticas⁷³.

FIGURA 1 - Normas de descrição arquivísticas



Fonte: Elaborado pelo autor

⁷³ Segundo o modelo conceitual de a *Norma Española Descripción Archivística* (2007), uma entidade arquivística é “qualquer objeto concreto ou abstrato que existe, existiu ou pode chegar a existir” e sobre o que se possa proporcionar informação (um documento, uma função, um produtor, uma norma, um lugar ou uma matéria).

As normas de descrição facilitam a identificação, a representação, a explicação e a gestão das diferentes entidades e suas inter-relações. Esse fato ratifica que a descrição não só se concentra no conteúdo dos documentos, mas também presta especial atenção a seu contexto de produção e, portanto, à elaboração normalizada de registros de autoridade. A normatização da descrição arquivística fez também repensar a ampliação do objeto da descrição. A partir de então, a descrição despende mais atenção à ação de explicar o documento e seu contexto de produção, para entender sua origem e procedência e, por isso, estabelece que nas descrições deve aparecer informação sobre o produtor.

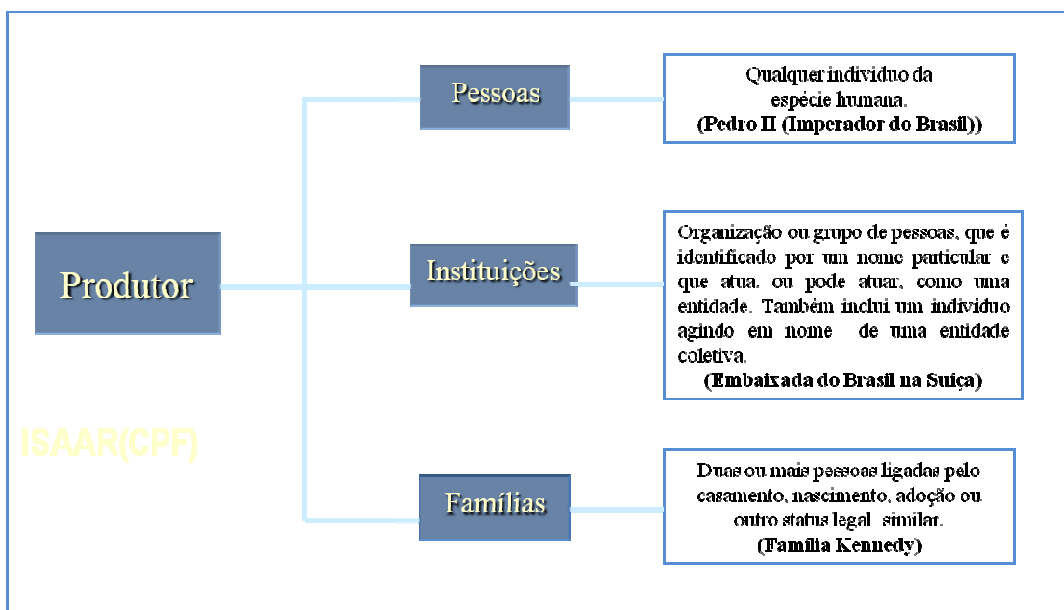
A maioria das normas publicadas desenvolvem estruturas de dados e conteúdos que facilitam a descrição da entidade documento de arquivo, as dezenove normas existentes (estrutura, conteúdos, apresentação e formatos), onze abordam pautas para a descrição dessa entidade e suas subentidades. Não obstante, o *International Council Of Archives* (ICA) e outros países elaboraram normas que permitem descrever, também, os produtores arquivísticos e, por conseguinte, criar registros de autoridade. Em breve, essas normas serão referidas.

Nos sistemas atuais de descrição arquivística é quase obrigatório que se descrevam os produtores. Tal processo requer extensa documentação e atualização contínua do contexto de produção e uso dos documentos, especialmente de sua procedência (ISAAR(CPF), 2004, p. 8), entendendo-se por procedência “relação existente entre os documentos de arquivo e as pessoas físicas ou jurídicas que os produziram, reuniram, conservaram e utilizaram no desenvolvimento de sua atividade pessoal ou institucional” (ISAAR(CPF), 2004, p.12).

Os produtores são qualquer entidade que cria, produz, gerencia e conserva os documentos no desenvolvimento de sua atividade pessoal ou institucional (ISAAR(CPF), 2004, p.12). Eles desempenham diversos papéis em relação aos documentos de arquivo, podendo ser autores, produtores, colecionadores e gestores.

Os produtores são uma entidade arquivística composta por três subentidades bem definidas: instituições, pessoas e família. À continuação demonstra-se, graficamente, em que consistem:

FIGURA 2 - Subentidades do produtor



Fonte: Elaborado pelo autor

O ICA, para a descrição dos produtores e a redação de registros de autoridades de arquivos, elabora uma norma, International Standard Archival Authority Record (Corporate bodies, Persons, Families) (ISAAR(CPF)), em 1995. A partir dessa norma, elaboraram-se outras que regulam a criação dos registros de autoridade. A saber: EE.UU publicam Describing Archives: A Content Standards (DACS), em 2004; Portugal, as Orientações para a Descrição Arquivística (ODA), em 2007; na Espanha, no ano 2000, a comunidade de Castilla e León publica o Manual de Descripción Multinivel (MDM) y Aragón, e em 2008, a Norma Aragonesa para la descripción de Autoridades de Archivos (ARANOR).

A norma internacional propõe uma estrutura de dados que deve ser contemplada para a elaboração de qualquer tipo de registro de autoridade nos arquivos. Essa estrutura tem sido aceita por os países (EE. UU, Portugal y Espanha) que desenvolveram normas para descrever o produtor. A tabela seguinte mostra as áreas e os elementos contemplados em ISAAR(CPF).

TABELA 8 - Estrutura de dados para a criação de registros de autoridade proposto em ISAAR(CPF)

ESTRUTURA DE DADOS PARA REGISTROS DE AUTORIDADE ISAAR(CPF)	
1. ÁREA DE IDENTIFICAÇÃO	3. ÁREA DE RELACIONAMENTOS
1.1 Tipo de entidade. 1.2 Forma(s) autorizada(s) do nome. 1.3 Formas paralelas do nome 1.4 Formas normalizadas do nome de acordo	3.1 Nome(s) / Identificadores das entidades coletivas, pessoas ou famílias relacionadas. 3.2 Categoria do relacionamento. 3.3 Descrição do relacionamento.

com outras regras. 1.5 Outras formas do nome. 1.6 Identificadores para entidades coletivas	3.4 Datas do relacionamento.
2. ÁREA DE DESCRIÇÃO	4. ÁREA DE CONTROLE
2.1 Datas de existência. 2.2 História. 2.3 Locais. 2.4 Status legal. 2.5 Funções, ocupações e atividades. 2.6 Mandatos / fontes de autoridade 2.7 Estrutura(s) interna(s) / Genealogia 2.8 Contexto geral	4.1 Identificador do registro de autoridade. 4.2 Identificadores da instituição. 4.3 Regras e/ou convenções. 4.4 Status de elaboração. 4.5 Nível de detalhamento. 4.6 Data de revisão, criação ou obsolescência. 4.7 Idioma(s) e sistema(s) de escrita. 4.8 Fontes. 4.9 Notas de manutenção.

Fonte: ISAAR(CPF)

Estes elementos são essenciais para a criação de registros de autoridade completos dos três tipos de produtores arquivísticos (instituição, pessoa ou família) e constituem um guia geral para os arquivistas. Embora todos os elementos apresentados sejam significativos para elaborar um registro de autoridade, o ICA outorga o caráter obrigatório a quatro elementos: Tipo de entidade; Forma autorizada do nome; Datas de existências; Identificador do registro de autoridade, necessários para estabelecer o intercâmbio de informação nos entornos eletrônicos e para descrever o produtor com um nível básico de informação.

O modelo internacional e os demais modelos foram criados com a finalidade de elaborar registros de autoridade de arquivos que proporcionem descrições de entidades (instituições, pessoas e famílias) associadas à produção e à gestão de arquivos. (ISAAR(CPF), 2004, p.8).

Mas, o que se entende em arquivologia por registro de autoridade? Um registro de autoridade, segundo a norma internacional, não é mais que uma “forma autorizada do nome, combinada com outros elementos informativos que permitem identificar e descrever a entidade mencionada e que podem também remeter a outros registros de autoridade relacionados” (ISAAR(CPF), 2004, p.12). Por sua vez, uma forma autorizada do nome é um ponto de acesso que identifica, de maneira unívoca, uma instituição, pessoa ou família e é estabelecido por um serviço de arquivo responsável (ISAAR(CPF), 2004, p.16).

A determinação do ponto de acesso é importante na conformação do registro de autoridade arquivístico, por determinar qual é a forma autorizada do nome da entidade que se descreve. Nas normas internacionais observa-se esse conceito de forma concisa; o *General International Standard Archival Description* (ISAD(G)) define os pontos de acesso como “os nomes, palavras chave, frases ou códigos que podem ser usados para buscar, identificar e

localizar uma descrição arquivísticas”, (ISAD(G), 2000, p.17); enquanto ISAAR(CPF) os define como “nome, termo, palavra chave, expressão ou código utilizado para identificar, procurar ou localizar descrições de documentos de arquivo”, (ISAAR(CPF) 2004, p.12).

Os profissionais do âmbito dos arquivos, apoiando-se nas experiências desenvolvidas nas bibliotecas, publicaram um conjunto de normas para a elaboração e a normatização de pontos de acesso de nomes de instituições, pessoas e famílias. É recomendável utilizar essas normas para a elaboração de listas de pontos de acesso nas instituições arquivísticas⁷⁴.

Para formalizar os pontos de acesso nos arquivos, é necessário usar as normas mencionadas e, além disso, ter em conta as seguintes questões:

- Usar várias fontes de informação para formalizar o ponto de acesso.
- Controlar sinonímias e homonímias.
- Usar qualificadores.

O processo pelo qual se normalizam os pontos de acesso denomina-se controle de autoridades; ou seja, o controle de termos normatizados que compreendem os nomes (de pessoas, entidades ou nomes geográficos) utilizados como pontos de acesso.

O controle de autoridades é um processo intelectual e de pesquisa pelo qual se determina a forma única de expressar um mesmo ponto de acesso e suas relações nos registros. Esse processo implica analisar todas as variantes possíveis com as que se pode denominar uma entidade arquivística, evitando as confusões dos homônimos e sinônimos.

O conjunto de registros é o que se denomina fichário de autoridades, que resulta ser uma lista de nomes (pessoas, entidades, congressos, famílias e lugares geográficos), de títulos uniformes (individuais ou coletivos) ou de matérias (palavras chave ou descritoras) (ISO 5123/3^a, 1981; UNE 50113/3a, 1991), (HERRERO, 1999, p. 121). Por sua vez, ISAAR(CPF) o define como um conjunto organizado de registros de autoridade, (ISAAR(CPF), 2004, p.12).

As descrições arquivísticas têm que representar e relacionar todas as entidades, portanto, a elaboração de um registro de autoridade (produtor) deve realizar-se em conjunto

74 As normas publicadas para a normatização de pontos de acesso vinculados a nomes de instituições, pessoas ou famílias são relacionadas a seguir: Describing Archives: A content standard. (DACS). EE.UU; Règles pour la Description des Documents d'Archives. (RAD2). Canada; Rules for the construction of personal, place and corporate names. (RCPPCN). Gran Bretaña; Orientações para a descrição arquivística. (ODA). Portugal; Manual de descripción multinivel: propuesta de adaptación de las normas internacionales de descripción archivística. (MDM). España Castilla y León; Norma para la elaboración de los puntos de acceso normalizados de instituciones, personas, familias, lugares y materia en el sistema de descripción archivística de los Archivos Estatales. (NEPAN). España.

com o fichário de descrições (documento). Para poder relacionar ambos os fichários, é necessário contemplar um elemento nas descrições que facilite esta inter-relação. ISAD(G) propõe o elemento *nome do produtor* para facilitar essas relações e os nexos entre a descrição dos documentos e sua procedência.

A interação que se produz entre o fichário de autoridades e o fichário de descrições é o que se denomina sistema de autoridades (ISAAR(CPF), 2004, p.14). Os sistemas de autoridades permitem a interação da informação contextual com a informação descritiva; sendo fundamentais na recuperação de informação do material arquivístico, de onde resulta indispensável a implementação de sistemas integrais de autoridades nas instituições arquivísticas.

Como se pode observar, a normatização dos registros de autoridade nos arquivos já é um feito. Vinculado ao processo de descrição de autoridade, os arquivistas devem ter presentes conceitos como registro de autoridade, pontos de acesso, controle de autoridades, fichários e sistemas de autoridades. Os arquivistas, nos últimos anos, têm dedicado tempo e esforço para apresentar as concepções teóricas da descrição e uma base normativa sólida para o trabalho nos arquivos.

A criação de registros de autoridade não é um processo exclusivo dos arquivos; as bibliotecas têm um longo e consistente trabalho vinculado a esse tema. Por isso, considera-se importante apresentar, a seguir, quais são as semelhanças e as variações entre os registros de autoridade nestes sistemas de informação.

4 REGISTROS DE AUTORIDADE EM BIBLIOTECAS E ARQUIVOS: SEMELHANÇAS E DIFERENÇAS

Como foi exposto anteriormente, a criação e a manutenção de registros de autoridade é uma prática que se desenvolve tanto em bibliotecas quanto em arquivos. Nestes sistemas de informação em que é preciso contar com sistemas de autoridades para consolidar a representação e recuperação da informação.

Com o propósito de mostrar as semelhanças e as diferenças entre os registros de autoridade das duas instituições de informação estudadas, elabora-se uma tabela comparativa que exemplifica os elementos de um registro de autoridade em bibliotecas em arquivos.

TABELA 9 - Registro de autoridade de instituições de informação

ELEMENTOS QUE INTEGRAM OS REGISTROS DE AUTORIDADES EM BIBLIOTECAS E EM ARQUIVOS.			
ELEMENTOS	B	A	

1	Cabeçalho Autorizado/ Forma autorizada do nome	X	X
2	Cabeçalho paralelo / Formas paralelas do nome	X	X
3	Formas normalizadas dos nomes de acordo com outras regras	-	X
4	Nota de informação	X	-
5	Traçado de referência de ver / Outras formas do nome	X	X
6	Traçado de referência de ver também / Outras formas do nome	X	X
7	Nota do catalogador/ Fontes	X	X
8	Número Internacional Normalizado de Dados de Autoridade / Identificador do registro de autoridade	X	X
9	Tipo de entidade/Nota de informação	X	X
10	Identificadores para entidades coletivas	-	X
11	Datas de existência.	-	X
12	História / Nota de Informação	X	X
13	Locais	-	X
14	Status legal	-	X
15	Funções, ocupações e atividades	-	X
16	Mandatos / fontes de autoridade	-	X
17	Estrutura(s) interna(s) / Genealogia	-	X
18	Contexto geral	-	X
19	Nome(s) / Identificadores das entidades coletivas, pessoas ou famílias relacionadas	X	X
20	Categoria do relacionamento	-	X
21	Descrição do relacionamento	-	X
22	Datas do relacionamento	-	X
23	Identificadores da instituição / Nome da agência bibliográfica nacional	X	X
24	Regras e/ou convenções / Regras de catalogação ou normas	X	X
25	Status de elaboração.	-	X
26	Nível de detalhamento	-	X
27	Data de revisão, criação, ou obsolescência / Data	X	X
28	Idioma(s) e sistema(s) de escrita	-	X
29	Notas de manutenção	-	X

Fonte: ISAAR(CPF) (ICA) y GARE (IFLA)

Como se pode observar na tabela anterior, foram identificados 29 elementos de informação que podem formar parte de um ou outro registro de autoridade. Desse total, 42% dos elementos são comuns aos dos registros (bibliotecas e arquivos). Os elementos iguais estão vinculados aos nomes autorizados das entidades, às formas paralelas, a outras formas de nomes, às fontes, aos códigos de registros de autoridade, às relações e aos dados referentes às instituições responsáveis pela elaboração e manutenção dos registros de autoridade.

Os registros de autoridade de arquivos estão compostos por um número maior de elementos (27). O ICA para elaboração desta estrutura de dados baseou-se, indiscutivelmente, nos estudos teóricos existentes no campo da biblioteconomia e nas normas das bibliotecas para a criação de registros. Como resultado, em um fichário de autoridade de arquivos está presente a maioria dos elementos propostos pelo IFLA. Nos arquivos, os registros de autoridades, além de controlar a forma do nome e sua identidade, há de se documentar a

informação sobre o produtor e o contexto de criação dos documentos. Essa é a razão pela qual os registros de autoridade de arquivos contêm mais elementos do que os de bibliotecas.

Em ambos os registros a informação está estruturada em áreas de informação mas não há correspondência entre elas (ver TAB. 9). Cada instituição cria áreas para estruturar seus registros de acordo com os objetivos próprios de cada sistema. Por exemplo, as bibliotecas brindam maior relevância às referências a outros tipos de cabeçalhos vinculados ao principal e, para isso, são criadas duas áreas de informação; os arquivos, porém, dispensam mais atenção a áreas vinculadas ao contexto e às relações entre os produtores.

TABELA 10 - Áreas de informação dos registros de autoridade em bibliotecas e em arquivos

ÁREAS DE INFORMAÇÃO	
BIBLIOTECAS	ARQUIVOS
Área de entrada autorizada.	Área de identificação.
Área de nota de informação.	Área de descrição.
Área de referência ver.	Área de relações.
Área de referência ver também.	Área de controle
Área de nota do catalogador.	
Área de fontes	

Fonte: Elaborado pelo autor

Os registros de autoridades em bibliotecas estão direcionados aos cabeçalhos de pessoas, entidades e obras/expressões, enquanto nos arquivos estão direcionados às entidades produtoras de documentos arquivísticos: instituições, pessoas e famílias. Como se pode observar, as criações intelectuais e artísticas são o elemento diferenciador entre ambas as instituições, pois nos arquivos esta entidade não existe.

Tanto nas bibliotecas quanto nos arquivos, para elaborar um registro de autoridade realiza-se controle de autoridade dos termos com os nomes utilizados como ponto de acesso.

Pesquisa de termos de autoridade em catálogos e bases de dados,

Criação da autoridade,

Estabelecimento de relações entre distintas autoridades,

Verificação e validação dos registros,

Manutenção da lista de autoridades, (HERRERO, 1999, p.130-132).

Outro aspecto determinante na criação de um registro nessas instituições é a criação do Cabeçalho autorizado/Forma autorizada do nome. Para normalizar a forma dos nomes realizam-se, também, operações similares:

- Usar várias fontes de informação.
- Usar qualificadores.

- Controlar as sinonímias.
- Controlar as homonímias.

O controle de autoridades e, por conseguinte, as normalizações dos pontos de acesso reduzem o tempo de trabalho e possibilitam o uso da informação em bibliotecas e arquivos, permitindo a integração de descrições de autoridade procedentes de diferentes instituições de informação em um mesmo sistema, facilitando a difusão e a disseminação.

Para finalizar, é conveniente mostrar um exemplo em que se expõe um registro de autoridade de uma mesma pessoa a partir de um enfoque biblioteconômico e arquivístico.

TABELA 1 - Registro de autoridade de Isabel I, rainha da Espanha

EXEMPLO DE REGISTRO DE AUTORIDADE		
ELEMENTOS	ISABEL I, RAINHA DA ESPANHA	
	BIBLIOTECA	ARQUIVO
Cabeçalho Autorizado/ Forma autorizada do nome	Isabel I, Rainha da Espanha	Isabel I (1451-1504, rainha de Castilla)
Nota de informação	-	-
Traçado de referência de ver / Outras formas do nome	Isabel, a Católica, Rainha da Espanha	Isabel I (1451-1504, rainha consorte de Aragón)
Traçado de referência de ver também / Outras formas do nome	Isabel I, Rainha de Castilla Isabel de Castilla Isabel de España, de William Th. Walsh, 1993. Isabel la Católica, de Tarsicio de Azcona, 1993.	-
Nota do catalogador/ Fontes	Dic. de Historia de España (Isabel I la Católica, Reina de Castilla y de Aragón (1451-1504); reinó de 1474-1504).	-
Número Internacional Normalizado de Dados de Autoridade / Identificador do registro de autoridade	XX1146476	ES28079MCU1576
Tipo de entidade/Nota de informação	-	Pessoa
Datas de existência.	-	22-04-1451 / 1504 Nasce em 22 de abril de 1451 em Madrigal das Altas Torres. Era a terceira filha do rei Juan II de Castilla, casado em segunda núpcia com dona Isabel de Portugal. Em 1468,
História / Nota de Informação	-	

Nome(s) / Identificadores das instituições, pessoas ou famílias relacionadas	Castilla (Reino). Rey (1474-1504: Isabel I). España. Rey (1474-1504: Fernando V e Isabel I). España -- Historia -- 1474-1516 (Reyes Católicos)	morto já seu irmão o infante Alfonso, firma-se o Pacto de los Toros de Guisando em que Enrique reconhece sua irmã Isabel como princesa de Asturias, deserdando por isso sua filha Juana. Em 1469 casa com Fernando de Aragón, filho de Juan II. Foi à morte de Enrique IV quando deflagrou uma guerra civil que enfrentou os partidários de Isabel e os de Juana. Os Tratados de Alcaçovas (setembro de 1479) põem fim à contenda e desde esse momento Isabel está firmemente assentada no trono.
Natureza da relação	-	Família (matrimônios)
Descrição da relação	-	Esposo
Nome(s) / Identificadores das instituições, pessoas ou famílias relacionadas	-	<u>Fernando II (1452-1516, rei de Aragón)/</u> ES28079MCU1946
Natureza da relação	-	Família (descendente de)
Descrição da relação	-	Filha e herdeira
Identificadores da instituição / Nome da agência bibliográfica nacional	BNE19900168656	Seção Nobreza do Arquivo Histórico Nacional. ES45168SNAHN
Estado de elaboração	-	Finalizado
Nível de detalhe	-	Básico
Língua(s) e escritura(s)	-	SPA

Fonte: Registro de Autoridade da Biblioteca Nacional de Espanha / Censo-Guia de Arquivos do Ministério de Cultura de Espanha.

Com suas semelhanças e suas diferenças, os registros de autoridades em ambos os sistemas de informação conquistam seus objetivos: facilitar a representação das autoridades, a precisão na recuperação, a cooperação entre as redes de bibliotecas e os arquivos e o acesso a toda a informação conservada nesses centros de memória coletiva. É importante destacar que

a existência de diretrizes em ambas as instituições para a criação dos registros permite a inter-relação entre diferentes contextos e o intercâmbio internacional de dados de autoridade nos ambientes eletrônicos.

5 CONSIDERAÇÕES FINAIS

A criação e manutenção de registros de autoridade é um processo que se realiza tanto em bibliotecas quanto em arquivos. Ambas as instituições, para a formação de fichários de autoridades, regem-se por modelos que estão normalizados pelas instituições internacionais e nacionais responsáveis pela atividade. Os bibliotecários e arquivistas realizam operações similares para a criação e formação dos registros de autoridade, os pontos de acesso e o controle normatizado de autoridades. Finalmente, pode-se concluir que os profissionais da Ciência da Informação deixaram de lado suas divergências e têm contribuído conjuntamente, sem esquecer suas particularidades, com o objetivo de obter integração e intercambiar, assim, os dados de autoridade da comunidade internacional.

REFERÊNCIAS

- ASSUMPCÃO, F. S.; SANTOS, P. L. V. A. da C. **A importância do controle de autoridade**: uma abordagem baseada nos objetivos e nas funções dos catálogos. In: ENCONTRO NACIONAL DE CATALOGADORES, 1., 2012, Rio de Janeiro; ENCONTRO DE ESTUDOS E PESQUISAS EM CATALOGAÇÃO, 3., 2012, Rio de Janeiro. Disponível em: <<http://eprints.rclis.org/18843/1/assumpcao-santos-a-importancia-do-controle-de-autoridade.pdf>>. Acesso em: 15 jul. 2014.
- BOERIS, C. **Algunas reflexiones sobre el control de autoridades en Argentina**. In: ENCUESTRO INTERNACIONAL Y NACIONAL DE CATALOGADORES, 7., 3., 2011, Buenos Aires. Anais... Buenos Aires: Biblioteca Nacional de la República Argentina. Disponível em: <<http://www.bn.gov.ar/descargas/catalogadores/encuentro2011/ponencia-boeris.pdf>>. Acesso em: 15 jul. 2014.
- CAMPOS, F. M. **Controle de autoridade**: novos contextos e soluções. Lisboa: Biblioteca Nacional, 2003. Disponível em: <<http://purl.pt/331/1/docs/comunicacao/12manha/controlodeautoridade.doc>>. Acesso em: 15 jul. 2014.
- COMISIÓN DE NORMAS ESPAÑOLAS DE DESCRIPCIÓN ARCHIVÍSTICA. Modelo conceptual de descripción archivística y requisitos de datos básicos de las descripciones de documentos de archivo, agentes y funciones. Madrid: CNEDA, 2012. Disponível em: <http://www.mcu.es/archivos/docs/NEDA_MCDA_P1_P2_20120618.pdf>. Acesso em: 02 jun. 2014.
- CONSEJO INTERNACIONAL DE ARCHIVOS. ISAD (G): **Norma internacional para la descripción archivística**. Adoptada por la comisión ad hoc sobre normas de descripción, Estocolmo, Suecia, 19-22 de septiembre de 1999. Madrid: Dirección General del Libro, Archivos y Bibliotecas, 2000.

CONSEJO INTERNACIONAL DE ARCHIVOS. **Norma internacional sobre los registros de autoridad de archivos relativos a instituciones, personas y familias**. 2. ed. Madrid: Dirección General del Libro, Archivos y Bibliotecas, Subdirección General de los Archivos Estatales, 2004.

DECLARAÇÃO de princípios internacionais de catalogação. [s.l.]: IFLA, 2009. Disponível em: <http://www.ifla.org/files/assets/cataloguing/icp/icp_2009-pt.pdf>. Acesso em: 15 jul. 2014.

HERRERO, C. El control de autoridades. **Anales de Documentación**, n. 2, p. 121-136, 1999. Disponível em: <<http://revistas.um.es/analesdoc/article/view/2621>>. Acesso em: 28 maio 2014.

MEY, E. S. A.; SILVEIRA, N. C. **Catalogação no plural**. Brasília: Briquet de Lemos, 2009.

OLIVEIRA, F. S. G. de. **O controle de autoridade na rede virtual de bibliotecas: Congresso Nacional (RVBI) e o modelo FRAD**. 2013. 121f. Monografia (Bacharelado em Biblioteconomia) – Faculdade de Ciência da Informação, Universidade de Brasília, Brasília, 2013.

PRAGMÁTICA NA ORGANIZAÇÃO DO CONHECIMENTO

PRAGMATIC IN KNOWLEDGE ORGANIZATION

Leila Cristina Weiss
Marisa Bräscher

Resumo: Caracteriza a abordagem semântica pragmática na organização do conhecimento com base na análise da literatura da área. Utiliza a técnica de análise de conteúdo para caracterizar a abordagem pragmática, identificando ideias compartilhadas pelos diferentes autores dos textos que compõem o corpus da pesquisa, formado por 79 trabalhos. Agrupa as ideias comuns e complementares em sete enunciados que destacam aspectos importantes do paradigma pragmático na área de organização do conhecimento. Dentre as características desse paradigma, destaca-se a importância de possibilitar uma comunicação efetiva na construção de representações da informação e do conhecimento, por meio da compreensão dos diferentes pontos de vista e não da escolha de um como sendo o melhor para representar uma realidade que é, por natureza, heterogênea. Conclui-se que as relações semânticas podem desempenhar um papel crucial nessa comunicação, ao guiar o usuário a partir do seu ponto de vista inicial até aqueles defendidos pela frente de pesquisa.

Palavras-chave: Pragmática. Organização do Conhecimento. Relações semânticas.

Abstract: It features the pragmatic semantic approach to knowledge organization based on analysis of previous researches. It uses the content analysis technique to characterize the pragmatic approach, identifies shared ideas by different authors of the texts that make up the research corpus, consisting of 79 works. It groups the common ideas and supplementary ideas in seven statements that highlight important aspects of the pragmatic paradigm in knowledge organization. Among the characteristics of this paradigm, it is emphasized, in the construction of representations of information and knowledge, the importance of enabling effective communication, by the comprehension of different points of view other than the choice of one as being the best to represent a reality that is by nature heterogeneous. It is concluded that the semantic relationships can play a crucial role in this communication, by guiding the user from his initial point of view to those advocated by the research front.

Keywords: Pragmatic. Knowledge Organization. Semantic relationships.

1 INTRODUÇÃO

Frequentemente os termos “pragmática” e “pragmatismo” são usados de forma equivalente. No entanto, assim como Marcondes (2000) consideramos que é preciso distinguir a *pragmática* enquanto um campo de estudos da linguagem e o *pragmatismo* enquanto corrente filosófica, ainda que uma filosofia da linguagem na linha da pragmática e o pragmatismo se aproximem em muitos aspectos.

O pragmatismo, como corrente filosófica, refere-se às concepções de filosofia que defendem, sobretudo, o primado da razão prática em relação à razão teórica. Teve origem no final do século XIX e desenvolveu-se ao longo do século XX, principalmente nos Estados Unidos. “Charles Sanders Peirce (1839-1914), William James (1842-1910) e posteriormente

John Dewey (1859-1952), são os principais representantes desse pensamento em suas várias vertentes.” (MARCONDES, 2000, p. 39).

Na linguística, a pragmática se caracteriza pelo estudo da linguagem em uso e faz parte do campo geral de estudos da linguagem, tradicionalmente dividido entre a pragmática, a semântica e a sintaxe. No entanto, conforme esclarece Marcondes (2000), existe ainda a concepção de pragmática na qual se valoriza a linguagem comum e o uso concreto da linguagem como a principal instância de investigação da linguagem. Nessa concepção considera-se que a constituição do significado linguístico se dá a partir de alguns aspectos: a interação entre falantes e ouvintes, o contexto de uso, os elementos socioculturais pressupostos pelo uso, e os objetivos, efeitos e consequências desses usos. “A pragmática não seria assim apenas um segmento dos estudos da linguagem, mas o seu campo privilegiado.” (MARCONDES, 2000, p. 40)

O pragmatismo como corrente filosófica influencia as pesquisas de diferentes áreas do conhecimento, inclusive a área de estudos da linguagem. Os pesquisadores em epistemologia definem essa influência como paradigma ou visão de mundo, que não é necessariamente apenas uma influência que ocorre posteriormente ao desenvolvimento de uma corrente filosófica, mas pode ser também um enquadramento ou compatibilidade com determinada corrente.

Quando alguns dos praticantes da ciência descobrem contradições internas e chegam à conclusão de que a sua maneira de ver o mundo não é adequada, percebem que o mundo deveria, ou ao menos poderia, ser olhado e investigado de outra maneira (KUHN, 2001).

Na área de estudos da linguagem encontramos estudos epistemológicos que consideram a semântica como a área principal, o significado o seu objeto de estudo, e a pragmática um paradigma em semântica. Hjørland (2007b) ressalta que Peregrin (2004) destaca dois paradigmas dominantes em semântica, o pragmático e o positivista. O paradigma semântico positivista seria a semântica formal, com regras de lógica para inferências, na qual não se considera o contexto para inferir o significado das expressões, levando-se em conta apenas o sentido denotativo, por que o considera a verdade. No paradigma pragmático em semântica se considera o contexto onde as expressões ocorrem para a inferência dos significados, aceitando-se, portanto, diferentes interpretações.

Peregrin (1999) trata especificamente sobre a pragmatização da semântica e afirma que Charles Morris e Rudolf Carnap estabeleceram uma divisão triádica para as teorias da linguagem: **sintaxe** trataria as relações entre as expressões; **semântica** abordaria as relações entre as expressões e o que elas representam (significados); **pragmática** examinaria as

relações entre as expressões e os seus usos. No entanto, assim como Peregrin (1999, 2004), entendemos que a semântica e a pragmática deveriam ser estudadas juntas, a semântica formal não deveria ser desprezada, mas entendida “como uma forma de destacar e materializar cada expressão contribuindo para as inferências de acordo com o contexto onde ocorrem” (PEREGRIN, 2004, p. 12, tradução nossa).

A pragmática alterou a maneira de estudar o significado. Ainda que o pragmatismo, como corrente filosófica, possa extrapolar os estudos da linguagem e do significado, nessa pesquisa aceitamos a pragmática como um dos paradigmas em semântica que pode contribuir com o estudo do significado na Organização do Conhecimento (OC) e na Recuperação da Informação (RI). Na literatura da área de Ciência da Informação (CI) é evidenciada a ligação da área de semântica com os campos de OC e RI. Existe um consenso entre os diferentes autores que a área de semântica pode fornecer subsídios teóricos para os campos de OC e RI. As teorias semânticas da vertente pragmática são apontadas em trabalhos como Blair (2003) e Hjørland (2007b) como uma opção frutífera para a OC. Mas, como efetivamente podemos caracterizar o pensamento dessa vertente nos estudos sobre OC? Como podemos pensar a elaboração de sistemas de organização do conhecimento (SOC) sob a perspectiva da abordagem pragmática?

A partir desses questionamentos, esta pesquisa tem por objetivo caracterizar a abordagem semântica pragmática na organização do conhecimento, com base na revisão de literatura da área de Ciência da Informação (CI). Aproximando as considerações encontradas no *corpus* de pesquisa, que foram majoritariamente de cunho teórico e sobre aspectos da CI no geral, dos aspectos envolvidos na OC. Neste trabalho sintetizamos os resultados obtidos por meio da análise de um corpus constituído de 79 textos, detalhado no item a seguir.

2 PROCEDIMENTOS METODOLÓGICOS

Esta pesquisa caracteriza-se como revisão de literatura e aplica técnicas da análise de conteúdo, definidas por Bardin (1979), para sistematização das principais ideias extraídas dos textos que compõem o corpus de pesquisa. O levantamento bibliográfico para a constituição do corpus de pesquisa foi realizado em fontes de informação especializadas da área, conforme consta no QUADRO 1.

QUADRO 1 – Levantamento bibliográfico

Fonte	Estratégia de busca (Campo e termo)		Documtnos recuperados	Documentos selecionados
BRAPCI	Geral - todos	Pragmática	28	12
	Geral -todos	Pragmatismo	9	4

LISA	Assunto principal	<i>pragmatic OR pragmatism</i>	29	11
ARIST	Verificação dos sumários		2	2
Repositórios ⁷⁵	pragmática OR pragmatismo		3 teses 3dissertações	3 teses 3dissertações
<i>Web of Science</i>	Hjørland		24	24
	Frohmann		6	6
Plataforma Lattes	Carlos Candido Almeida		8	8
	Gustavo Silva Saldanha		1	1
	Luciana de Souza Gracioso		1	1
	Maria Nélide Gonzalez de Gomez		4	4
Total 79 documentos				

Fonte: autoria própria.

Os seis autores pesquisados na Plataforma *Lattes* e na *Web of Science* foram identificados como importantes para o entendimento da concepção do paradigma pragmático em OC. Por esse motivo, julgamos relevante um levantamento mais exaustivo da sua produção científica.

Os documentos selecionados foram analisados com o intuito de identificar ideias comuns ou complementares que caracterizam o paradigma pragmático. Essas ideias foram agrupadas, criando-se assim categorias. Essa análise foi realizada em três fases, definidas por Bardin (1979), pré-análise, exploração do material e tratamento dos resultados (inferência e interpretação). Quanto ao tipo de análise empregada, a categorial, a autora descreve que

No conjunto das técnicas da análise de conteúdo, a análise por categorias é de citar em primeiro lugar: cronologicamente é a mais antiga; na prática é a mais utilizada. Funciona por operação de desmembramento do texto em unidades, em categorias segundo reagrupamentos analógicos. Entre as diferentes possibilidades de categorização, a investigação dos temas, ou análise temática, é rápida e eficaz na condição de se aplicar a discursos diretos (significações manifestas) e simples. (BARDIN, 1979, p.153).

Para Bardin (1979, p.117) “a categorização é uma operação de classificação de elementos constitutivos de um conjunto, por diferenciação e, seguidamente, por reagrupamento segundo o gênero (analogia), com os critérios previamente definidos.”

Esses critérios podem ser semânticos (categorias temáticas), sintático (os verbos, os adjetivos), dentre outros. Além disso, a categorização pode empregar dois tipos de processos, que são tidos como inversos. Esses processos recebem os nomes de “por caixas”, no qual é fornecido o sistema de categorias e repartem-se da melhor maneira possível os elementos, à medida que vão sendo encontrados, e “por milha”, no qual o sistema de categorias não é

⁷⁵ repositórios institucionais ou nos catálogos das bibliotecas de universidades brasileiras que possuem programa de pós-graduação em Ciência da Informação, de acordo com a classificação da Capes.

fornecido antes, resultando da classificação analógica e progressiva dos elementos. O título conceitual de cada categoria, somente é definido no final da operação. (BARDIN, 1979)

Nessa pesquisa o critério empregado para a formação das categorias foi o semântico. Os temas foram identificados na medida em que os trabalhos que fazem parte do corpus da pesquisa foram analisados. Ou seja, as categorias ou classes não foram previamente definidas. O processo empregado então foi o “por milha”, como denominado por Bardin (1979). Após a categorização, foram propostos sete enunciados que, em nossa análise, sintetizam as principais características do paradigma pragmático nos estudos da OC, e nomeiam as subseções do item a seguir em que os resultados são apresentados e discutidos.

3 RESULTADOS E DISCUSSÃO

Nessa seção apresentam-se as características da abordagem pragmática na OC. Os títulos das subseções sintetizam o tema de cada uma. As discussões apresentadas foram construídas de modo a argumentar e justificar o agrupamento das ideias, identificando os pontos em comum ou que se complementam.

3.1 A informação é um fenômeno construído

Redón Rojas (1996) considera que a informação é construída e não simplesmente inventada (subjetiva) ou descoberta (objetiva). A informação não é uma simples invenção, pois para sua elaboração se toma certa matéria prima (elementos objetivos que se encontram nas coisas), e nem é, tampouco, algo pronto e objetivo que só precisa ser descoberto, pois é necessária a ação do homem. Os homens transformam e são transformados pelo mundo informacional em um processo de evolução constante. O conceito de informação pragmática, em que a dialética do sujeito com o mundo que o rodeia é levada em consideração, deveria ser um dos pilares dos estudos em biblioteconomia, segundo o autor.

Esclarecemos que a ideia de informação como fenômeno construído, o produto de uma síntese de elementos objetivos e subjetivos, de propriedades reais dos objetos com a atividade do homem, é uma visão da CI. Hjørland (2003, p. 93) descreve que os pesquisadores de CI “não são *experts* na interpretação de informações das estrelas, mas são, na sua maior parte, especialistas na informação documentada por astrônomos”. O autor apresenta esse exemplo, no qual a informação é entendida em um sentido mais amplo do que normalmente está implícito na CI, para fundamentar a seguinte afirmação:

O foco especial da CI é o conhecimento documentado produzido pelos seres humanos em algum tipo de documento de uso potencial para outros seres humanos. A luz das estrelas não é informação para a comunidade da CI, mas informações astronômicas como as produzidas e utilizadas por astrônomos

são. Esta distinção pode parecer sutil, mas é importante, a fim de construir uma base teórica sólida para a OC (HJØRLAND, 2003, p. 93).

Nesse sentido, Hjørland(2007 a) considera apropriado falar-se em signos ao discutir aspectos da OC na CI e considera a teoria semiótica como um quadro teórico de referência mais apropriado para os estudos da informação do que a teoria de Shannon e Weaver, por exemplo. Para o autor teorias objetivistas e universalistas da informação podem ter um apelo muito mais forte do que pontos de vistas teóricos que fazem informação, significado e decisões dependentes do contexto. No entanto, segundo o mesmo autor, uma consideração superficial da natureza da informação nos deixaria sem uma fundamentação teórica adequada.

Na tradição semiótica pragmática desenvolvida pelo filósofo e lógico estadunidense Charles Sanders Peirce (1839-1914) os conceitos são entendidos como signos. De acordo com Thellefsen e Thellefsen (2004, p. 180) “isso nos permite afirmar que o vasto conteúdo de conhecimento tácito em um domínio de conhecimento expressa padrões estáveis de significação, que é o resultado de processos sígnicos.” Esses processos sígnicos (semiose) posteriormente tornam-se hábitos de interpretações.

De acordo com a nossa compreensão do conhecimento, a interpretação dos signos em um contexto social não depende apenas do indivíduo intérprete, mas está ancorada no contexto social. Não se pode interpretar os signos de acordo com caprichos subjetivos ou preferências. O domínio do conhecimento coloca restrições sobre as nossas interpretações (THELLEFSEN; THELLEFSEN, 2004, p. 180).

Para Hjørland (2009, p. 1530), “conceitos tem sido entendidos como significados socialmente negociados”. Por se tratar de um processo social, essa negociação e, conseqüentemente, a formação dos conceitos, ocorrem em contextos múltiplos e heterogêneos, que devem ser levados em conta, uma vez que deles dependem. A OC e a elaboração de SOC deveriam, portanto, analisar os contextos de uso para identificar conceitos e possíveis relacionamentos. Assim continuamos a discussão sobre esse tema na seção a seguir.

3.2 Informação e conceitos ocorrem em contextos múltiplos e heterogêneos

Vários estudos da CI, como os bibliométricos e de OC, são desenvolvidos no contexto de determinado domínio do conhecimento. Isso se deve, dentre outros fatores, à necessidade de minimizar as diferenças conceituais. No entanto, diferenças conceituais e concepções concorrentes podem ocorrer mesmo no interior de determinada área do conhecimento.

A análise de domínio, desenvolvida por Birger Hjørland e Hanne Albrechtsen, assume que diferentes abordagens (ou paradigmas) existem em todos os domínios do conhecimento. De acordo com Hjørland (2008b), esses paradigmas devem ser identificados no processo de

OC. No entanto, o autor esclarece que podem existir abordagens ou paradigmas não representativos, que não são apropriados para a OC. Dessa forma, qualquer SOC acaba apresentando um tipo de inclinação, ou *bias*, para os pontos de vistas mais representativos da área.

Outro aspecto da inevitável inclinação nos SOC diz respeito à linguagem em si. Como destaca Gonzales de Gomez (2001), Wittgenstein colocou uma espécie de sombra sobre as concepções universalistas da linguagem, uma vez que, na sua visão, não existe nenhuma linguagem capaz de sintetizar outras linguagens; e os paradigmas de Kuhn fazem o mesmo sobre os territórios discursivos da ciência. Para Kuhn (1975, p. 437 *apud* GONZALEZ DE GOMEZ, 2001, p. 8) “as linguagens seccionam o mundo de maneiras diferentes e não dispomos de nenhum meio sublinguístico neutro de informar”.

Para Alvarenga (1998, p. 3) “a construção do conhecimento deve ser vista como um processo eivado de interesses pessoais e subjetividade, devido aos imperativos que governam a conduta social implícita na construção não somente da ciência, mas dos saberes em geral” e afirma:

O conhecimento científico repousaria sobre um suporte institucional, sendo reforçado e acompanhado por outros estratos e práticas sociais, tais como a política, pedagogia, o sistema de comunicação do conhecimento, incluindo os sistemas de editoração e bibliotecas. [...] a biblioteca seria vista como uma instituição para arranjar textos, estabelecendo-se como um componente da legitimação de uma ordem particular do discurso (ALVARENGA, 1998, p. 3).

Os desafios para este arranjo de textos, que é um componente de legitimação, são analisados por Hjørland (2012), tanto no nível teórico quanto no prático. No nível prático, as bibliotecas estariam cada vez mais dispensando a classificação de livros. No plano teórico, muitos pesquisadores, gestores e usuários acreditam que a atividade de “classificação” não valeria mais a pena devido ao esforço necessário, pois consideram que os motores de busca podem fornecer metadados igualmente úteis para a RI. No nível prático, Hjørland(2012) considera que a questão básica na classificação é a escolha de critérios para decidir se algo deve ser classificado como X.

Os motores de busca podem ser calibrados de forma a proporcionar diferentes resultados ou rankings. A fim de fazer essa calibração (ou simplesmente para avaliar os sistemas), é preciso ter algum tipo de classificação do que deve ser encontrado. Até agora, no campo da CI, temos utilizado principalmente as avaliações de relevância baseadas na "relevância do usuário" (ver HJØRLAND, 2010) (HJØRLAND, 2012, p. 311).

Em Hjørland (2010 a) encontramos argumentação acerca da seleção e determinação da relevância.

A seleção do sistema (ou seja, do programador) não é, portanto, "perfeita" ou "objetiva"; é uma escolha feita entre muitas possíveis e que pode ser mais ou menos inteligente ou adequada para diferentes fins. Portanto, é subjetiva (é difícil entender porque tantas pessoas afirmam o contrário). Sua subjetividade é determinada pelas escolhas do programador, incluindo opções de pontos de acesso de assunto, de ponderação e utilização de estruturas de ligação. Além disso, e talvez o mais importante: a escolha dos métodos de avaliação (HJØRLAND, 2010a, p. 218).

O autor prossegue a sua argumentação e descreve a diferença entre os conceitos de “objetivo” e “subjetivo”. “Uma instrução ou representação é subjetiva se remete às opiniões, crenças e sentimentos de convicção desse ou daquele indivíduo. É objetivo se é independente da opinião das pessoas, crenças e pontos de vista” (HJØRLAND, 2010a, p.218).

Hjørland (2013, p. 11) considera que os SOC não podem ser objetivos e assim devem ser baseados em algum tipo de subjetividade. Esta subjetividade deve, no entanto, ser derivada de opiniões coletivas nas comunidades discursivas e não de estudos de indivíduos ou a partir do estudo abstrato da mente.

Frohmann (2008b), considera que a subjetividade é o tema central para estudos sobre a ética da informação. Para desenvolver sua argumentação o autor apresenta a contribuição de diversos autores (Froehlich, Hauptman, Foucault, Deleuze e Hacking). Sobre o pensamento de Gilles Deleuze, o autor afirma que,

Num espírito foucaultiano, ele argumenta que a ética da informação precisa ser genealógica, porque é preciso olhar para as possibilidades morais e éticas inseridas nos regimes de informação que habitamos.[...] a individualidade e subjetividades possíveis em qualquer momento histórico estão vinculadas com os sistemas sócio-técnicos que as tornam possíveis. A ética da informação deve levar em conta as possibilidades da ação ética onde as forças de estabilização, codificação, territorialização, e dominação no trabalho de construir e configurar os modos de subjetivação possam sofrer negociação, oposição, resistência, e transmutação. Na medida em que a ética está preocupada com a subjetividade, o problema não está localizado em debates sobre como determinados sujeitos morais devem agir, mas na interação de poder entre as forças de dominação e as possibilidades de liberdade na formação de sujeitos. Precisamos de uma ética da informação que reconheça como os processos e tecnologias de informação estão envolvidos no desenvolvimento dos indivíduos (FROHMANN, 2008b, p. 273).

Para Frohmann (2008b, p.275) “um estudo abrangente da importância de Foucault para a ética da informação ainda precisa ser feito.” Assim, procura mostrar que uma ética da liberdade de Foucault oferece uma crítica à ética da informação *self-centered*. O autor explica a diferença entre a ética da liberdade de Foucault e a ética *self-centered* de Froehlich e Hauptman, afirmando que, no último caso,

as subjetividades são muito duráveis e muito estáveis. Apesar de serem componentes emergentes de *assemblages* morais, suas configurações são

tomadas como dadas de uma vez por todas, sem a possibilidade de suas contingências ou o fracasso de suas *assemblages* jamais serem recebidas. (FROHMANN, 2008b, p. 276)

Assemblages, segundo Frohmann (2008b, p.275) se referem às “configurações em que muitos elementos heterogêneos (físicos, conceituais, discursivos, tecnológicos, institucionais) estão alinhados para produzir o todo. Os elementos específicos que compõem *assemblages* variam caso a caso [...]”.

o conceito de *assemblage*, especialmente elaborado por Latour, ajuda-nos a situar os documentos e a documentação em associações complexas de entidades amplamente heterogêneas. Também nos redireciona para a ética e a política dos estudos de documentação que são sensíveis às forças de associação ou reunião, que são tão importantes para a análise de Latour. Ele nos mostra como a documentação pode ser uma questão ética e política importante para a concepção das formas de vida coletiva. (FROHMANN, 2008a, p. 179)

Essa e outras questões são abordadas pelo autor, que conclui o artigo afirmando que

Foucault, Deleuze e Hacking nos apresentam uma forma de pensamento ético que revela problemas e questões em ética da informação veladas a partir da perspectiva de uma ética da informação *self-centered*. Seus trabalhos também nos confrontam com alternativas, não só dos modos de dominar a subjetivação dos sistemas de informação, mas também de como lidar com questões morais daí provenientes. Mas também sugere uma crítica: à medida em que uma ética da informação *self-centered* contribui para a estabilização dos modos de subjetivação e *assemblages* da qual emergem, é cúmplice da reificação das forças de dominação das quais a sua ética quer nos libertar. (FROHMANN, 2008b, p. 276).

Hjørland (1998) se refere ao pensamento de Foucault ao afirmar que na OC é importante um tipo de análise relacionada ao “construtivismo social”, uma espécie de escavação de camada após camada (o que Michel Foucault chamou de “A Arqueologia do Saber”). Para Hjørland (2003, p. 94) a visão do construtivismo social “está relacionada com a visão pragmática [...]. Muitas vezes, porém, o construtivismo social e pragmatismo se opõem ao tipo de realismo, como realismo científico.” Mas a própria questão do realismo na visão pragmática não é um consenso, conforme afirmam Hjørland (2003) e Sudin e Johannisson (2004).

Se uma realidade, que é exterior à mente humana, existe ou não, não podemos afirmar. O fato é que, na visão pragmática, a interpretação humana dos fenômenos é levada em conta para se entender a construção do conhecimento. O pensamento de Foucault e a linha de estudos do construtivismo social são considerados úteis para a OC, apesar de o construtivismo social ser considerado antirrealista, e o antirrealismo não aceito na visão pragmática. Hjørland (2004) argumenta a favor do realismo na CI, e inclui a abordagem pragmática como integrante dessa perspectiva.

Frohmann (2008b, p.275) citando Latour (2005) nos ajuda a compreender como o construtivismo social pode ser compatível com o realismo.

O absurdo de supor que mostrar que algo é construído para diminuir a sua realidade ou para mostrar que é uma farsa é execrado por Latour em seu livro de 2005, no qual ele diz: “em todos os domínios, dizer que algo é construído tem sido sempre associado à valorização de sua robustez, qualidade, estilo, durabilidade, valor, etc. Tanto que ninguém se deu ao trabalho de dizer que um arranha-céu, uma usina nuclear, uma escultura, ou um automóvel são construídos. Isso é obvio demais para ser apontado”. [...] “quando dizemos que um fato é construído, significa simplesmente que representa a realidade objetiva e sólida mediante a mobilização de várias entidades cuja *assemblage* pode falhar.”

O fato do desenvolvimento do conhecimento depender da ação humana e cada indivíduo ser influenciado por alguns pressupostos, concepções e influências teóricas possui, em nossa opinião, uma relação de causalidade com a temática abordada na seção a seguir.

3.3 A qualidade da informação não é constante

A questão da qualidade da informação foi estudada por Assis e Moura (2011). As autoras observaram na literatura que as pesquisas não consideram a qualidade da informação como um processo, o que deveria ser reconsiderado uma vez que “nem informação nem qualidade são fenômenos constantes por que se modificam o tempo todo” (BRIER, 2006 *apud* ASSIS; MOURA, 2011, p. 16). Considera-se, assim, que os modelos de qualidade da informação compõem um recorte arbitrário e operacionalizável de um universo multidimensional e complexo de significações. São criados para contextos e propósitos específicos e desse modo, propensos a sofrerem desatualizações (ASSIS; MOURA, 2011).

De acordo com Hjørland (2000)

A produção e avaliação do conhecimento não podem ser feitas apenas por princípios empiristas ou racionalistas, mas por uma combinação, associando o conhecimento histórico da origem das teorias e considerando os objetivos e valores humanos. Conhecimento se torna mais contextualizado quando o analisamos sob a perspectiva dos documentos e de seu conteúdo. Um documento tem uma história, tem um ou mais autores ou produtores, tem uma conexão com outros documentos, e assim por diante (HJØRLAND, 2000, p. 34 e 35)⁷⁶.

Hjørland (2011) argumenta que para avaliar uma fonte de informação, um verbete da Wikipédia, por exemplo, é necessário relacionar o conteúdo dessa fonte com a interpretação do estado do conhecimento na frente de pesquisa, que normalmente é mais desenvolvido. Para

⁷⁶ “O ponto de vista do conhecimento como fatos isolados ou ideias está relacionado com o empirismo e o racionalismo e enquanto epistemologia pragmática, considera o conhecimento como um conjunto de teorias que cumprem algum propósito” (HJØRLAND, 2000, p.34).

o mesmo autor, as fontes de pesquisa devem ser avaliadas em relação à forma como abordam as controvérsias e diferentes pontos de vista.

Esses diferentes pontos de vista geram diferenças conceituais que são tradicionalmente vistas como problemáticas para a RI e os SOC seriam responsáveis por minimizá-las. Para isso, a compreensão de como ocorre o desenvolvimento da linguagem e dos significados pode ser útil, como passamos a discorrer.

3.4 A prática é importante para a construção do significado

Frohmann (2004) considera que a investigação de Wittgenstein sobre “o que é o significado?” é importante para estudos da informação. Descreve alguns exemplos simples extraídos do livro ‘Investigações Filosóficas’, que considera contribuir para deslocar nossa atenção das imagens mentalistas de significado para as práticas da linguagem.

Com base em Wittgenstein, Hjørland (1998, p.21) afirma que

significados são produzidos por nossas práticas sociais. Uma consequência da prática social é o desenvolvimento da comunicação, do comportamento verbal e não-verbal e de conceitos. Os significados são produzidos primeiramente “fora da mente” e são então, por meio da linguagem, transferidos para as mentes individuais. A partir dessa perspectiva, a questão central da semântica não está mais relacionada aos objetos ou mentes individuais, mas às culturas, às subculturas, à divisão social do trabalho, ao discurso das comunidades, às disciplinas científicas, e assim por diante.

Assim, o autor considera que as teorias mais adequadas sobre semântica seriam a sociocognitiva e a sociolinguística e não apenas a cognitiva e a linguística. No contexto mais amplo, da teoria sociocognitiva e da sociolinguística, o autor aponta para as tradições pragmáticas em semântica que precederam a teoria dos jogos de linguagem. Mais especificamente o pensamento de John Dewey e de Peirce.

Para John Dewey línguas são apenas um meio de comunicação de significado. Comunicação não-verbal, arte e objetos são todos expressivos; eles carregam significado, e podem ser considerados como uma espécie de linguagem. Cada arte tem o seu próprio meio e este meio é especialmente equipado para um tipo de comunicação. As necessidades da vida diária atribuíram mais importância ao discurso como forma de comunicação, em detrimento de outras. Diferentes culturas e necessidades humanas desenvolvem suportes especiais para comunicar significados. Para mim, esse ponto de vista parece intimamente relacionados com a teoria de Wittgenstein de ‘jogos de linguagem’ (HJØRLAND, 1998, p.22).

Sobre o pensamento de Peirce, Hjørland (1998 a, p.21) afirma que ele “descobriu que a teoria pragmática do significado é “futurista” interpretando o significado do ponto de vista de como a determinação do sentido pode contribuir para o cumprimento das metas.”

O pragmatismo e o realismo de Charles Sanders Peirce, que integram a sua teoria semiótica, por levar em conta a condição de referencialidade é considerado por Almeida

(2012b) uma opção teórica na OC para substituir o extremo relativismo. Almeida (2012a, p. 53) apresenta uma revisão da máxima do pragmatismo,

A fim de determinar o significado de uma concepção intelectual, deve-se considerar quais consequências práticas poderiam conceberivelmente resultar, necessariamente, da verdade dessa concepção; e a soma destas consequências constituirá todo o significado da concepção.

Mas Almeida (2012b, p. 207) também esclarece que,

a ação por si só não faz um conceito adequado, uma vez que o seu propósito depende da concepção criada antecipadamente e que dá sentido e referência para a ação, e não de outro modo. Como método, o pragmatismo visa averiguar o verdadeiro significado de qualquer conceito, doutrina, proposição, palavra, ou outro sinal.

Assim, certos tipos de concepções intelectuais não possuem significação pragmática. Almeida (2012a) apresenta o exemplo do conceito teológico de transubstanciação – a transformação do vinho em sangue e da hóstia em carne – que, ao aplicar-se o teste pragmático mostra-se não ter consequências práticas concebíveis. “Um conceito científico ou símbolo, por outro lado, respeitando as indicações do objeto no processo de produção do interpretante, admite um exame pragmático” (ALMEIDA, 2012a, p.53).

O conceito é produto da convenção e da formação de hábitos, mais do que resultado de simples qualidades, mesmo que estas caracterizem e individualizem o conceito em uma rede de conceitos. Deste modo, o conceito é um símbolo, resultante da união entre signo e objeto. Os conceitos dependem da formação de hábitos, os quais estabelecem as regras de associação e de aproximação do *representamen*⁷⁷ com o objeto (ALMEIDA 2012 a, p.52).

O papel das regras na teoria de Peirce, que dependem do hábito e aproximam o signo do objeto, é semelhante na teoria dos jogos de linguagem de Wittgenstein. Novelino (1998, p. 142) afirma que “o uso da linguagem, no sentido dado por Wittgenstein, implica no domínio e uso de regras, pois a aplicação correta de um termo significa que se age de acordo com as regras estabelecidas pelo contexto de sua aplicação.” As expressões linguísticas teriam significado apenas nos diferentes jogos de linguagem que são formações complexas de linguagem e ação.

O grau das diferenças conceituais entre os jogos de linguagem variam, e de forma geral sempre existem conceitos compartilhados entre os diferentes jogos de linguagem, o que possibilita a interação entre esses. Com a interação é possível validar pontos de vista novos, diferentes daqueles que já se tinham e já se confiava, e diminuir progressivamente as diferenças conceituais. Para elaboração de SOC devemos encontrar mecanismos que

⁷⁷ Sinônimo de signo.

possibilitem compreender as práticas discursivas e identificar a evolução dos conceitos em determinado domínio do conhecimento.

3.5 A interação é importante para a validação

Ao voltar sua atenção para as ações de busca de informação em ambientes que possibilitam a interação, Gracioso (2008) investiga o processo de aceitação e validação de propostas informacionais. A autora buscou na teoria da ação comunicativa de Habermas, bem como em pesquisas baseadas nessa teoria, a resposta sobre como os sujeitos aceitam e validam a informação. Uma das pesquisas analisadas pela autora é a de Gerald Benôit que desenvolveu testes em sistemas informacionais e concluiu que a aproximação dos critérios de significação construídos entre “humanos” é difícil de ser apropriada (*a priori*) completamente.

Guasque (2008), ao buscar identificar a relação entre o pensamento reflexivo proposto por John Dewey e as competências empregadas na busca e no uso da informação, verificou que

- Ao se depararem com conhecimentos novos e pontos de vista diferenciados, pesquisadores em formação articulam e aplicam os critérios de autoridade e coerência da abordagem para avaliar a informação e suas fontes, analisando a estrutura e lógica dos argumentos ou métodos. [...]
- Pesquisadores em formação validam a sua compreensão e interpretação por meio de conversas com outros colegas, especialmente os do grupo de estudo e os professores (GUASQUE, 2008, p. 202).

Esses resultados empíricos, ao nosso ver, reforçam os argumentos de Gracioso (2008), quando afirma ser possível, por um processo interativo, atingir critérios de validação e estabelecer um entendimento mútuo sobre as ações proferidas, o que não necessariamente pode resultar em acordos.

Sobre a questão de acordos Guedes (2010), em trabalho que busca insumos teóricos nos pressupostos do pensamento dialógico de Mikhail Bakhtin⁷⁸ (1895-1975) para elucidar a natureza das ações interdiscursivas recorrentes na indexação social, afirma que

O diálogo deve ser visto como uma das manifestações do dialogismo. Fiorin (2006b, p. 24) explica que o vocábulo “diálogo”, portador do significado de “solução de conflito”, “entendimento”, “busca de acordo”, entre outros, “pode levar a pensar que Bakhtin é o filósofo da grande conciliação entre homens. Ao contrário, as relações dialógicas podem ser contratuais ou polêmicas, de divergências ou de convergências, de aceitação ou recusa [...]”. A ligação entre os discursos, isto é, o movimento interdiscursivo

⁷⁸ “o principal tema presente nos textos do Círculo de Bakhtin era o estudo da linguagem no processo de interação social, o que mais tarde seria difundido pelo Círculo pelo termo dialogismo” (GUEDES, 2010, p. 51).

pregado pelo dialogismo, é sempre assimétrico, heterogêneo e conflituoso (GUEDES, 2010, p. 61).

Sendo assim, pode-se concluir que por meio da interação, pontos de vista são contrapostos e podem ou não ser validados, aceitos como verdadeiros. A abordagem pragmática na OC considera que se existirem pontos de vista diferentes, esses devem ser apresentados aos usuários. Na seção a seguir descrevemos como esses pontos de vista devem ser identificados para então serem apresentados aos usuários.

3.6 A bibliografia é uma fonte para identificar pontos de vista distintos e construir linguagens de comunicação

A abordagem pragmática em CI, conforme descrevemos ao longo desse trabalho, assume que diferentes abordagens (ou paradigmas) existem em todos os domínios do conhecimento. Assim, esses devem ser ativamente pesquisados na OC, pois “tais pontos de vista diferentes podem ser explícitos ou implícitos e, se forem implícitos, podem ser descobertos pela análise teórica e filosófica” (HJØRLAND, 2007b, p.374).

Um item central para essa análise seria a bibliografia da área. Hjørland (2007b) retoma o princípio da garantia literária, cunhado em 1911 pelo britânico Edward Wyndham Hulme, o qual diz respeito à identificação e validação dos elementos dos SOC por meio da análise da bibliografia.

Em qualquer bibliografia encontramos maneiras distintas de definir conceitos e determinar relações semânticas. Garantia literária não significa identificar apenas um texto do qual as relações semânticas podem ser inferidas. A tarefa é negociar entre diferentes critérios e selecionar os que tem o maior grau de autoridade cognitiva ou são considerados melhores em relação ao objetivo do SOC (HJØRLAND, 2007b, p.388 e 389).

Se percebe uma estreita ligação entre a importância dada aos documentos, à bibliografia, para os estudos informacionais e a forma como o sujeito é visto, como um ser social e não individual. Nesse sentido, destacamos Frohmann (2009), Hjørland (2007b), Almeida, Cedón e Pinheiro (2012) e Campos e Venâncio (2006).

Com a análise dos documentos e dos usos da linguagem seria possível entender a dinâmica social para a formação do conhecimento. Saldanha (2008, p.22) considera que

Ao ir contra o horizonte de uma racionalidade e de um cientificismo estáticos, essencialistas – a representação -, o pragmatismo informacional procura discutir que nenhuma classificação de mundo dura mais que sua linguagem de uso – e que, mesmo entre os becos e travessas que jogam com suas palavras, a dinâmica deste uso transforma permanentemente a estrutura das formas e dos significados. Não há nada definitivo na esfera das relações sociais. Tudo está sob a dinâmica permanente dos contextos com os quais os jogos de linguagem se constituem, se transformam e desaparecem.

Assim, o objetivo de representar o conhecimento nos SOC deveria ser substituído pelo objetivo de comunicar, sem forçar uma estabilidade que não existe. A representação do conhecimento é, tradicionalmente, vista como o objetivo da OC. Um SOC visa representar o conhecimento de determinado domínio, para organizar a informação e aperfeiçoar o processo de RI.

Para Novelino (1998) linguagens formalizadas são imprescindíveis para ações de transferência da informação, ou comunicação formal da informação, no entanto, as linguagens formalizadas construídas para esse fim são linguagens de representação e essas refletem uma grande preocupação com aspectos representacionais da linguagem quando deveriam refletir uma preocupação maior com o aspecto comunicacional. Sobre as diferenças entre as linguagens de representação e as de transferência de informação estaria o fato de que as de representação consideram exclusivamente a essência de cada informação analisada, enquanto as centradas na comunicação, as de transferência como a autora propõe, consideram os contextos de produção e de uso da informação (NOVELINO, 1998).

Na abordagem pragmática encontramos oposição a alguns aspectos tradicionalmente aceitos na RI e OC, tais como necessidade de informação, e as medidas de precisão e revocação, conforme discutimos a seguir.

3.7 Guiar o usuário para uma escolha informada

Um dos aspectos que nos chama a atenção ao analisar o pensamento dos autores que tratam sobre a pragmática na CI, é a crítica aos atuais modelos de recuperação da informação. Blair (2003) considera que

pode ser extremamente difícil conceber sistemas de informação radicalmente diferentes ou melhorados, porque estamos praticamente bloqueados na maneira de pensar sobre a recuperação da informação, que se materializa pelos sistemas existentes. (BLAIR, 2003, p.13e14, tradução nossa).

O processo de RI, muitas vezes, é visto como aquele no qual o usuário tem algo em mente, que seria a necessidade de informação, que é traduzida em uma consulta de pesquisa. No entanto, com base nas afirmações de Wittgenstein, o autor explica que a forma como se pensa a necessidade de informação é condicionada pela linguagem de recuperação disponível. Na medida em que essa linguagem é limitada, assim será o pensamento sobre o que se quer, sobre a necessidade de informação (BLAIR, 2003).

A avaliação dos sistemas de recuperação de informação (SRI) leva em conta, principalmente e talvez unicamente, os índices de precisão e revocação. Esses índices são

calculados a partir do julgamento de relevância, e a relevância entendida tradicionalmente como o que o usuário já teria em mente, aquilo que o usuário quer encontrar com a busca.

Pensar na diferença entre querer e precisar pode ajudar a entender essa questão. Aquilo que o usuário quer, que tem relação com o conceito tradicional de relevância, pode não ser o que ele precisa. O precisar tem uma ligação direta a aspectos funcionais e práticos enquanto querer é algo mais subjetivo. Assim, o precisar nos parece ser compatível com a definição de relevância, de Hjørland e Christensen (2002, p.964), “Algo (A) é relevante para uma tarefa (T) se aumentar a probabilidade de realizar o objetivo (G), que está implícito em T”.

Frohmann (1990, p. 98) também aborda esse tema e apresenta a seguinte questão: “A recuperação de textos deve satisfazer uma necessidade ou satisfazer um desejo?”.

Desejos são explicitamente reconhecidos e admitidos; eles refletem os objetivos dos agentes, os propósitos e intenções. Nem todas as necessidades são conhecidas, [...] Por exemplo, nem todos sabem o que precisam para evitar AIDS, e nem todo mundo deseja o que precisa. A identificação das necessidades depende de uma concepção da natureza humana e do mundo social; desejos podem ser identificados por meio de um questionário. Se apenas a satisfação do desejo for considerada como finalidade da recuperação de textos, a maioria das regras de indexação para as práticas de recuperação servirão à forma de organização social predominante. Entre as regras de indexação importantes para a satisfação do desejo no capitalismo de consumo, por exemplo, estão aquelas que representam eficientemente bens para o consumo. Por outro lado, se a recuperação objetiva satisfazer as necessidades, então as regras para a sua prática podem ser inconsistentes com os objetivos da ordem social dominante e até antagonistas a eles (FROHMANN, 1990, p. 98).

Essas e outras questões exigiriam uma análise bem mais ampla, ou uma análise política como o autor sugere. Em todo caso, o que essa questão deixa claro é que as “diferentes concepções sobre o papel social da recuperação de textos determinarão os tipos de regras de indexação que construímos” (FROHMANN, 1990, p. 98).

Basear-se prioritariamente nos “desejos” e àquilo o que o usuário já tem em mente pode ser limitativo e ocultar pontos de vista importantes presentes na literatura. De acordo com Gracioso (2010b, p. 141) “é importante que pesquisadores (buscadores de informações) tenham a dimensão da indeterminação da relação entre o que eles descreverão e o que, diante do que foi recuperado, terão de discriminar”. Para isso, conforme afirma Hjørland (2007b, p. 389) “talvez a tarefa mais importante do profissional da informação seja tornar os diferentes interesses e paradigmas visíveis para que o usuário possa fazer uma escolha informado”. Assim, talvez o usuário possa perceber que aquilo que ele queria não era o que ele efetivamente precisava.

Hjørland (2003) explica a influência dos diferentes paradigmas no estabelecimento das relações semânticas por meio da descrição da evolução do conhecimento científico. Apresenta como exemplo, dentre outros, a classificação dos animais. As baleias vivem na água e podem ser classificadas como animais aquáticos, as baleias também são mamíferos e não peixes. A classificação exige que se encontrem as propriedades semelhantes entre os itens a serem classificados para, então, agrupá-los. Essas propriedades semelhantes também podem estabelecer outras relações entre os itens. Hjørland (2003, p. 102) aponta que “a história de todas as ciências naturais revela que a descoberta de que certas entidades que, em princípio pensava-se partilharem determinadas propriedades, poderiam pertencer a diferentes espécies.”

Assim, o autor afirma que

essa realidade mais profunda de espécies e conceitos que a ciência descobre tem implicações importantes para a metodologia de OC. Em primeiro lugar ela desafia muitas abordagens orientadas ao usuário e empiristas. Na medida em que esta visão é correta, as relações entre dois conceitos são, portanto, em relação aos sistemas teóricos (ou paradigmas) em que estão incorporadas. [...] Podemos concluir que as unidades básicas na OC, as relações semânticas entre dois conceitos, devem ser relativas à perspectiva e à teoria da qual elas são consideradas. Devido a este fato, OC não pode ser feita apenas a partir de combinações sucessivas de elementos mas deve refletir perspectivas e teorias mais amplas (HJØRLAND, 2003, p. 102 e 103).

A maneira como o autor entende os conceitos, que junto com as relações semânticas formam os SOC, vai ao encontro dessa afirmação. De acordo com Hjørland (2009) conceitos, pelo menos para fins de OC, são entendidos como significados socialmente negociados, e uma justificativa para esse entendimento, que nos parece oposta ao entendimento do conceito como unidade do pensamento, seria o fato de que

Se o conhecimento é definido como crença verdadeira justificada (como na tradição platônica), então o conhecimento real é difícil ou impossível de identificar e classificar. É mais proveitoso falar de reivindicações de conhecimento, e não do conhecimento em si. Falar em reivindicações de conhecimento sobre as coisas representadas na literatura e as coisas a serem classificadas é uma maneira mais cuidadosa de falar, e não há perda real por esse modo de dizer (HJØRLAND, 2003, p. 100).

Ao nosso ver, isso também justifica o fato de as relações semânticas serem consideradas pelo autor como a unidade básica na OC. Pois na negociação dos significados se formam relações semânticas, ainda que a negociação não termine em um acordo e não chegue a um conceito unívoco, devido a pontos de vistas distintos, as relações semânticas estarão lá. Se forem devidamente identificadas, ajudarão a comunicar ou mostrar para o usuário o contexto, ou os contextos, nos quais o conceito está inserido.

4 CONSIDERAÇÕES FINAIS

Neste trabalho procuramos apresentar as principais características da abordagem pragmática na OC, campo no qual o estudo do significado, ou semântica, é de suma importância. O que, de certa forma, nos faz entender a pragmática como um paradigma em semântica, conforme sugere Peregrin (2004; 1999), visão que também é aceita por Birger Hjørland.

A partir dos resultados obtidos com essa pesquisa podemos dizer que na abordagem pragmática a informação é vista como um fenômeno que é construído. Essa construção se dá em circunstâncias diversas e envolve diferentes atores que podem possuir pontos de vista e necessidades distintas, que se dão por diversos aspectos contextuais envolvidos na produção e uso da informação e do conhecimento. Assim, a avaliação da qualidade da informação deve levar em conta estes pontos de vista distintos. Tanto a informação, quanto a sua qualidade não são vistos como fenômenos constantes e avaliar uma fonte de informação a partir de premissas de verdade ou falsidade é visto como perigoso, pois se deve avaliá-la em relação à forma como ela aborda as controvérsias e diferentes pontos de vista. Assim como fontes de informação tais como as enciclopédias, os SOC também devem apresentar os diferentes pontos de vista encontrados na literatura de determinada área do conhecimento.

Ao adotar princípios da abordagem pragmática, os SOC podem contribuir para uma comunicação efetiva, ao possibilitar a compreensão dos diferentes pontos de vista presentes numa comunidade discursiva e não optar por um como sendo o melhor para representar uma realidade que é, por natureza, heterogênea. As relações semânticas poderiam desempenhar papel crucial nessa comunicação, ao permitir que o usuário possa partir do seu ponto de vista inicial e ir até aqueles defendidos pela frente de pesquisa sobre o tema.

Os SOC devem acompanhar a evolução dos saberes e fazeres, refletida também nos usos da linguagem, que estariam presentes na literatura dos diferentes domínios. Nessa perspectiva, poderiam estar mais próximos dos estudos da Sociolinguística ou da Socioterminologia no estudo dos contextos de uso, nos quais, de acordo com a visão da abordagem pragmática, encontraríamos os conceitos e relacionamentos que comporiam a espinha dorsal dos SOC.

REFERÊNCIAS

ALMEIDA, C. C. Conceito como signo: elemento semiótico para análise e mediação da informação. **Scire**, Zaragoza, v. 18, p. 49-56, 2012a.

ALMEIDA, C. C. The Methodological Influence of Peirce's Pragmatism on Knowledge Organization. **Knowledge Organization**, v. 39, p. 204-215, 2012b.

ALMEIDA, M. B.; CEDON, B.V.; PINHEIRO, M. M.K. Princípios metodológicos para a caracterização da dimensão pragmática de documentos no desenvolvimento de ontologias biomédicas. **Inf. & Soc.:Est.**, João Pessoa, v.22, n.1, p. 105-117, jan./abr. 2012.

ALVARENGA, Lidia . Bibliometria e arqueologia do saber de Michel Foucault: traços de identidade teórico-metodológica. **Ciência da Informação**, Brasília, v. 27, n. 3, p. 253-261, 1998.

ASSIS, J.; MOURA, M. A. A qualidade da informação na *web*: uma abordagem semiótica. **Inf. Inf.**, Londrina, v. 16 n. 3, p. 96 – 117, jan./ jun. 2011.

BARDIN, L. *Análise de conteúdo*. Lisboa: Edições 70, 1979. 225p.

BLAIR, D. C. Information Retrieval and the Philosophy of language. **ARIST**, v. 37, p. 2-50, 2003.

CAMPOS, L.; VENÂNCIO, L. O objeto de estudo da ciência da informação: a morte do indivíduo. **Inf. Inf.**, Londrina, v. 11, n. 1, jan. /jun. 2006.

FROHMANN, B. Documentary ethics, ontology, and politic. **Arch Sci**, v. 8, 2008a.

FROHMANN, B. Documentation Redux: Prolegomenon to (Another) Philosophy of Informatio. **Library Trends**, v. 52, n. 3, 2004.

FROHMANN, B. Revisiting “what is a document?” **Journal of Documentation**, v. 65, n. 2, 2009.

FROHMANN, B. Rules of indexing: a critique of mentalism in information retrieval theory. **Journal of Documentation**, v. 46, n. 2, 1990.

FROHMANN, B. Subjectivity and Information Ethics. **JASIST**, v. 59, n. 2, 2008b.

GONZALEZ DE GOMEZ, M. N. Para uma reflexão epistemológica acerca da Ciência da Informação. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 6, n. 1, p. 5-18, jan./jun. 2001.

GRACIOSO, L. S. **Filosofia da linguagem e Ciência da Informação**: jogos de linguagem e ação comunicativa no contexto das ações de informação em tecnologias virtuais. 2008. 176f. Tese (Doutorado) – Universidade Federal Fluminense/IBICT.

GRACIOSO, L. S. Parâmetros teóricos para elaboração de instrumentos pragmáticos de representação e organização da informação na Web: considerações preliminares sobre uma possível proposta metodológica. **InCID: R. Ci. Inf. e Doc.**, Ribeirão Preto, v. 1, n.1, p. 138-158, 2010b.

GUASGUE, K. C. G. D. **O pensamento reflexivo na busca e no uso da informação na comunicação científica**. 2008. 240f. Tese (Doutorado) Programa de Pós-graduação em ciência da Informação, Universidade de Brasília, 2008.

GUEDES, R. M. **A abordagem dialógica na indexação social**. 2010. 186f. Dissertação (Mestrado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2010.

HJØRLAND, B. Arguments for philosophical realism in Library and Information Science. **Library Trends**, v. 52, n. 3, 2004.

- HJØRLAND, B. Concept theory. **JASIST**, v. 60, n. 8, p. 1519-1536, 2009.
- HJØRLAND, B. Deliberate bias in Knowledge Organization? **Advances in Knowledge Organization**, v. 11, p. 256-260, 2008.
- HJØRLAND, B. Documents, Memory Institutions and Information Science. **Journal of Documentation**, v. 56, n. 1, January 2000.
- HJØRLAND, B. Evaluation of an information source illustrated by a case study: effect of screening for breast cancer. **JASIST**, v. 62, n. 10, 2011.
- HJØRLAND, B. Fundamentals of Knowledge Organization. **Knowledge Organization**, v. 30, n. 2, p.87-111, 2003.
- HJØRLAND, B. Information retrieval, text composition, and semantics. **Knowledge Organization**, v. 25, n. 1-2, 1998.
- HJØRLAND, B. Information: Objective or Subjective/Situational? **JASIST**, v. 58, n. 10, 2007a.
- HJØRLAND, B. Is classification necessary after Google? **Journal of Documentation**, v. 6, n. 3, 2012.
- HJØRLAND, B. Semantic and Knowledge organization. **ARIST**, v. 41, p. 367-405, 2007b.
- HJØRLAND, B. The Foundation of the concept of relevance. **JASIST**, v. 61, n. 2, 2010a.
- HJØRLAND, B. User-based and Cognitive Approaches to Knowledge Organization: A Theoretical Analysis of the Research Literature. **Knowledge Organization**, v. 40, n. 1, 2013.
- HJØRLAND, B. What is Knowledge Organization (KO)? **Knowledge Organization**, v. 35, n. 2, p. 86-101, 2008b.
- HJØRLAND, B.; CHRISTENSEN, F. S. Work Tasks and Socio-Cognitive Relevance: a specific example. **JASIST**, v. 53, n. 11, p.960-965, 2002.
- KUHN, T. S. **A estrutura das revoluções científicas**. 6. ed. São Paulo: Perspectiva, 2001.
- MARCONDES, D. Desfazendo mitos sobre a pragmática. **ALCEU**, Rio de Janeiro, v.1, n.1, p. 38-46, 2000.
- NOVELINO, M. S. F. A linguagem como meio de representação ou de comunicação da informação. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 3, n. 2, p. 137 - 146, jul./dez.1998.
- PEREGRIN, J. Pragmatism und Semantik. In: Fuhrmann, A.; OLSSON, E. J. **Pragmatisch denken**. p. 89-108. Frankfurt am Main, Germany: Ontos. 2004.
- PEREGRIN, J. Pragmatization of Semantics. In: TURNER, K. **The semantics/Pragmatics Interface from Different Points of View**. Elsevier, Amsterdam. 1999. p. 419-442.
- RENDÓN ROJAS, M. A. Hacia um nuevo paradigma em bibliotecologia. **Transinformação**, Campinas, v. 8, n. 3, p. 17-31, set./dez. 1996.

SALDANHA, G. S. Viagem aos becos e travessas da tradição pragmática da Ciência da Informação: uma leitura em diálogo com Wittgenstein. 2008. 302f. Dissertação (Mestrado) – Universidade Federal de Minas Gerais, 2008.

SUNDIN, O.; JOHANNISSON, J. Pragmatism, neo-pragmatism and sociocultural theory: Communicative participation as a perspective in LIS. **Journal of Documentation**, v.61, n. 1, 2005.

THELLEFSEN, T. L; THELLEFSEN, M. M. Pragmatic semiotics and knowledge organization. **Knowledge Organization**, v. 31, n. 3, p. 177-187, 2004.

**DIMENSÕES SUBJETIVAS NO TRATAMENTO DA INFORMAÇÃO: ANALISANDO
O PROCESSO DE INDEXAÇÃO E IMPLICAÇÕES SUBJACENTES NA
RECUPERAÇÃO DA INFORMAÇÃO**

**SUBJECTIVE DIMENSIONS IN THE TREATMENT OF INFORMATION:
ANALYZING THE INDEXING PROCESS AND UNDERLYING IMPLICATIONS
TO INFORMATION RETRIEVAL**

Eliane Pawlowski de Oliveira Araujo
Claudio Paixão Anastácio de Paula

Resumo: A explosão informacional tem contribuído para alterar o cenário e os comportamentos em relação ao trato com a informação, principalmente a partir do início do Século XXI. Gradualmente, a preocupação com a oferta de meios para que o usuário possa recuperar informações que atendam demandas eficazes e particularizadamente vai se impondo ao propósito mais generalizado de priorizar a mera disseminação da informação. Um dos pontos chave para que o processo de recuperação da informação ocorra de forma adequada encontra-se na etapa de tratamento informacional, processo que compreende, dentre várias tarefas, a indexação. Esta atividade apresenta o desafio de atribuir termos que representem fielmente o objeto analisado. A qualidade da indexação está condicionada à qualificação do indexador e sua “imparcialidade” no processo. Entretanto, sabe-se ser quase inevitável que alguns elementos de julgamento subjetivo afetem a realização dessa atividade. O presente trabalho pretende apresentar o resultado parcial de estudo realizado em Bibliotecas Universitárias no qual se pretendeu identificar como as perspectivas individuais permeiam o processo de tratamento informacional. Para atendimento ao objetivo proposto – de verificação da influência da subjetividade na atividade de indexação – a pesquisa teve como eixo norteador a Abordagem Clínica da Informação, perspectiva que apresenta a possibilidade de investigar o comportamento informacional considerando a influência de elementos culturais, simbólicos, cognitivos e afetivos, bem como fatores psicodinâmicos – conscientes e inconscientes. Como métodos constituintes dessa abordagem destacam-se a entrevista, a análise de tarefa, a técnica do incidente crítico e o protocolo verbal. Os resultados obtidos possibilitaram compreender como os aspectos subjetivos se integram às competências individuais para influenciar o tratamento informacional repercutindo na recuperação de informação pelos usuários. As constatações possibilitadas pela pesquisa demonstram a necessidade de caminhar na busca pelo estabelecimento de diretrizes mais precisas de modo a possibilitar o estabelecimento de parâmetros mais eficazes para orientar a execução dessa atividade.

Palavras-chave: Recuperação da Informação. Tratamento da informação. Análise de assunto. Indexação. Catalogação.

Abstract: The informational boom has contributed to change the scenario and the behavior in relation to how the information is treated, especially from the beginning of the 21st century. Gradually the concern regarding the offer of means for the user to retrieve information that meets demands efficiently and specifically imposes itself to the purpose of prioritizing only the dissemination of information. One of the key aspects for the information retrieval process to occur adequately is the treatment of the information. This process comprehends, among various tasks, the activity of indexing. This activity presents the challenge of attributing terms which represent the analyzed object faithfully. The quality of the process of indexing is conditioned by how qualified the indexer is, and his or her impartiality in the process. However, it is known that it is inevitable that some elements of subjective judgment affect

this activity. This article aims to present the partial results of a study carried out in university libraries, with the objective of identifying how the individual perspectives permeate the treatment of information. To meet the proposed objective – to check the influence of the subjectivity in the indexing process – this research had as a guiding principle, the Clinical Approach to Information, a perspective that presents the possibility of investigating the information behavior considering the influence of cultural elements, symbolic, cognitive and affective as well as psychodynamic factors – conscious and unconscious. As methods constituents of this approach, it is included the interview, the task analyses, the clinical incident technique and verbal protocol. The results obtained enabled the understanding of how subjective aspects integrate themselves to individual competences to influence the process of treatment of information, having an impact on information retrieval by users in a digital environment.

Keywords: Information retrieval. Treatment of information. Subject analysis. Indexing. Cataloging.

1 INTRODUÇÃO

O consumo de informação pelos americanos em 2008, segundo a pesquisa “*How Much Information 2009?*”, foi de cerca de 1.3 trilhões de horas de informação, o que corresponde a uma média de 34 gigabytes diários por pessoa considerando 20 diferentes fontes de informação, desde jornais e livros até as tecnologias mais modernas. Esse cenário demonstra como têm se tornado complexa as atividades relacionadas ao tratamento da informação, especialmente no aspecto referente à recuperação: como tratar a informação de modo a possibilitar a recuperação exata daquilo que se deseja nesta imensidão informacional?

Segundo Garcia (2007) a recuperação da informação é uma preocupação antiga. A autora cita como exemplo algumas iniciativas, como o “*mundaneum*” idealizado por Paul Otlet no Século XIX, que tinha como objetivo tornar acessível o conhecimento àqueles que dele necessitavam, bem como o “*memex*”, desenvolvido por Vannevar Bush na década de 1940, que visava facilitar a recuperação da informação.

A visão integrada e indissociável entre tratamento e recuperação da informação pode ser vista nos primórdios da Ciência da Informação. Conforme mencionado por Borko (1968)

Ciência da Informação é a disciplina que investiga as propriedades e o comportamento da informação, as forças que regem o fluxo da informação e os meios de processamento da informação para acessibilidade e usabilidade ótimas. Está relacionada com o corpo de conhecimento que abrange a origem, coleta, organização, armazenamento, recuperação, interpretação, transmissão, transformação e utilização da informação. Isto inclui a investigação das representações da informação nos sistemas naturais e artificiais, o uso de código para a transmissão eficiente de mensagens e o estudo dos dispositivos e técnicas de processamento de informação tais como computadores (BORKO, 1968, p.3).

No âmbito das Bibliotecas a recuperação da informação é um dos aspectos mais importantes decorrentes do tratamento da informação. Segundo Araújo (2013) “tratar” a informação implica a execução de várias atividades, dentre elas a indexação, todas visando possibilitar ao usuário a recuperação da informação desejada com a maior acurácia, efetividade, precisão e rapidez possíveis.

Entretanto, segundo Kobashi (1994), apesar da existência de metodologias que orientam a prática da análise de assunto na indexação existe um vazio teórico e metodológico nas várias regras de elaboração de informações documentárias, pois estas regras não apresentam indicações objetivas sobre o modo de obter os produtos desejados. Seu êxito, segundo a autora, se deve ao bom senso e experiência do indexador, fato que leva a atentar para a influência dos aspectos subjetivos na realização no processo de indexação.

Entender como a subjetividade se integra aos comportamentos vinculados ao tratamento da informação pode fornecer indícios de como minimizar a interferência desse aspecto no processo favorecendo que a recuperação possa ocorrer da forma mais eficaz possível. Para adentrar nesse universo realizou-se uma pesquisa com bibliotecários que atuam em Bibliotecas Universitárias visando obter, pela via simbólica, percepções sobre os aspectos intrínsecos presentes nesta atividade.

Importante ressaltar que a ação de descrever um documento de acordo com seu assunto – definida pelo UNISIST (1981) como indexação – no contexto das bibliotecas possui terminologia e conceitos próprios, sendo considerada como uma das atividades inerentes ao processo de catalogação. Alguns estudos como o de Fujita (2003), Silva e Fujita (2004), Lancaster (2004) e Fujita *et al* (2010) apresentam diferentes perspectivas teóricas e metodológicas sobre a indexação e a catalogação. No presente trabalho, as diferenças terminológicas dessas atividades não foram abordadas, visto ter-se analisado o ponto de equivalência entre ambas que é a atividade de análise de assunto, tarefa que busca representar documentos por meio de termos visando sua posterior recuperação, procedimento inerente ao exercício do bibliotecário na qualidade de catalogador. Conforme afirmam Silva e Fujita (2004, p.142) “Apesar das divergências sobre semelhanças e diferenças entre os termos, a indexação alfabética de assuntos e a catalogação de assuntos são equivalentes porque são resultados de um mesmo processo: a análise de assunto”.

2 CONTEXTO TEÓRICO

Para atendimento ao objetivo proposto – de verificação da influência da subjetividade no processo de tratamento da informação – a pesquisa teve como eixo norteador a Abordagem

Clínica da Informação, perspectiva proposta por Paula (2011) que apresenta a possibilidade de investigar o comportamento informacional considerando a influência de elementos culturais, simbólicos, cognitivos e afetivos, bem como fatores psicodinâmicos – conscientes e inconscientes. O autor sugere a expressão "Abordagem Clínica da Informação" para designar uma perspectiva de trabalho inspirada na designação francesa *approche clinique* que tem por característica um olhar profundo do fenômeno da informação utilizando-se de uma perspectiva clínica (sem o viés psicopatológico) para atingir níveis de análise não usuais nos estudos comportamentais e cognitivistas tradicionais.

Ainda segundo Paula (2012), este método de abordagem consiste em “investigar o objeto sobre o qual se põe um problema, inserindo as informações coletadas na dinâmica particular desse objeto, reconhecendo e determinando certos estados, padrões, movimentos e alterações”. Considera-se que essa ação irá permitir descrever fenômenos, tecer diagnósticos, prognósticos ou prescrever intervenções. A conversão de um estudo de caso tradicional em um exercício clínico se dá pela adoção da postura de uma análise profunda do caso, na qual os sujeitos do estudo são compreendidos em suas interações com o contexto que os rodeia e com seus elementos intrínsecos. Dessa forma, chega-se a uma compreensão da sua dinâmica, da origem da sua condição atual e seu ciclo vital. No entender do autor, a adoção desta prática possibilita o “sair de cena” de uma postura mais funcionalista da relação com a informação e adentrar ao palco das ações numa busca intensa pelos “comos” e os “porquês” das ações, consideradas subjetivas e dotadas de significados.

Um dos métodos utilizados na pesquisa, compondo esta abordagem, tem como fundamento os aspectos simbólicos e o imaginário. Segundo Gilbert Durand (1997) o imaginário é considerado como o alicerce sobre o qual se constroem as concepções de homem e de mundo. Ao considerá-lo como algo que não é uma vaga abstração, pois segue regras estruturais com vistas a uma hermenêutica, o autor reabilita a dimensão dos arquétipos entendidos num conceito amplo como estruturas inatas que servem de matriz ou uma forma “imaterial” à qual os fenômenos psíquicos tendem a se moldar. Os estudos deste autor sobre o imaginário, segundo Oliveira e Maia (2008, p.1), procuraram colocar “a imagem, a imaginação e o imaginário no cenário dos estudos acadêmicos” e se consolidaram em uma teoria que

se organiza sob o método da convergência, isto é, os símbolos se (re)agrupam em torno de núcleos organizadores, as constelações, as quais são estruturadas por isomorfismos, que dizem respeito à polarização das imagens; indica que há estreita relação entre os gestos do corpo e as representações simbólicas. Os símbolos constelam porque são desenvolvidos

de um mesmo tema arquetípico, porque são variações sobre um arquétipo. (CEMIM *et al*, 2001).

A teoria de Gilbert Durand foi sistematizada pelo psicólogo Yves Durand (1988) que selecionou nove elementos para servirem de estímulo à elaboração de um desenho, uma narrativa e um questionário, compondo o Teste Arquetípico dos Nove Elementos – AT-9. O Teste, conforme explicitado por Cemin *et al* (2001), permite entender como o sujeito age frente suas angústias e enfrentamentos mais elementares. Segundo Estrada (2002, p.28), os elementos do AT-9 têm por finalidade provocar a “questão do tempo e da morte, com a finalidade de se encontrar um meio de resolver a angústia original”. A intenção, de acordo com Paula (2012), é utilizar estímulos arquetípicos “que têm o papel de colocar o problema trabalhado numa perspectiva de tempo, ameaça e finitude” para construir modos de enfrentamento de um problema. No contexto da pesquisa, o “problema” em questão é a decisão sobre a atribuição de termos na execução da atividade de análise de assunto.

O contexto de realização da pesquisa, ou seja, a atividade de análise de assunto, constitui-se como a etapa do tratamento temático da informação em que um documento é analisado visando a extração de conceitos que possam traduzir sua essência. A relevância desta atividade reside, dentre outros, no fato de que a correta seleção de palavras facilita o atendimento, por parte dos leitores, de sua demanda informacional em sistemas informatizados.

Naves (1996) considera a análise de assunto como a operação base para os procedimentos de recuperação da informação e ampara-se em Harris (1970) ao afirmar que a falsa ideia de que esta constitui-se uma área aparentemente simples “ocorre por absoluto desconhecimento da complexidade do processo que exige esforços [...] no sentido de seguir uma metodologia adequada para obter resultados satisfatórios” (NAVES, 1996, p.217). Além do problema da terminologia, a autora destaca a influência direta da pessoa que executa a atividade, pois “não há dúvidas de que o indexador interpõe suas próprias ideias e preconceitos na sua atuação de intermediário entre autores e usuários” (NAVES, 1996, p.221).

Ferneda (2003) destaca que a análise de assunto pode envolver a interpretação do documento com a finalidade de agregar temas que não estão diretamente explicitados em sua superfície textual visando torná-lo visível aos usuários de um sistema de informação. O que torna esse processo complexo, segundo o autor, é que a necessidade de informação do usuário será representada por meio de uma expressão de busca e predizer as palavras que foram usadas para representar os documentos pelo indexador – e que terá como resultado a

satisfação das necessidades do usuário – tem sido um desafio considerando a quantidade de documentos disponibilizados nos sistemas de informação. Para Ferneda (2003) não é suficiente predizer um ou mais termos utilizados para indexar os documentos mas, também, é necessário evitar a recuperação de documentos não relevantes minimizando o esforço em verificar a relevância dos mesmos. O autor, em suas conclusões, ressalta que os elementos envolvidos no processo de recuperação da informação são tipicamente linguísticos e que uma interpretação correta desses elementos refletirá positivamente na qualidade dos resultados em um processo de recuperação de informação.

3 DA PESQUISA

A pesquisa foi conduzida em uma biblioteca especializada integrante de um Sistema de Bibliotecas de uma instituição de ensino superior. O domínio em análise, donde foram selecionados os sujeitos para participação da pesquisa, é uma das bibliotecas que possui um dos maiores acervos bibliográficos do Sistema. Esta ambiência mostrou-se adequada aos objetivos da pesquisa uma vez que, nas bibliotecas com maiores acervos, o volume de entrada de exemplares, e conseqüente tratamento de informação para compor o catálogo bibliográfico, também são maiores. Os sujeitos da pesquisa foram três bibliotecários, denominados neste trabalho como S1, S2 e S3, cujo critério de seleção foi a experiência na atividade de catalogação. O quantitativo de sujeitos da amostra foi definido considerando o caráter da pesquisa que permitiu utilizar uma amostragem não probabilística, critério referenciado por Oliveira (2001) quando se trata de uma população homogênea, sendo possível a escolha de *experts* nas amostras definidas por intenção ou julgamento.

Dentre os métodos empíricos utilizados destacam-se a entrevista semiestruturada e a atividade de análise de tarefa. Um dos objetivos da entrevista foi conhecer a trajetória profissional dos sujeitos da pesquisa. Nesta etapa foram entremeados aspectos simbólicos e afetivos visando reunir elementos para compor também um cenário de sentimentos, percepções e construções afetivas relacionadas ao exercício profissional de cada pesquisado. Também foi incorporada a Técnica do Incidente Crítico (TIC) que “consiste em um conjunto de procedimentos para a coleta de observações diretas do comportamento humano, de modo a facilitar sua utilização potencial na solução de problemas práticos e no desenvolvimento de amplos princípios psicológicos” (FLANAGAN, 1973, p.99). Um incidente é definido por Flanagan (1973, p.100) como “qualquer atividade humana observável que seja suficientemente completa em si mesma para permitir inferências e previsões a respeito da pessoa que executa o ato”. Segundo o autor, para ser considerado crítico, “um incidente deve

ocorrer em uma situação na qual o propósito ou intenção do ato pareça bastante claro ao observador e suas conseqüências sejam suficientemente definidas para deixar poucas dúvidas sobre os seus efeitos”. O incidente analisado nesta pesquisa foi relacionado a uma situação de tomada de decisão na execução da atividade de análise de assunto que o entrevistado tenha considerado relevante destacar.

Na análise de tarefa realizada objetivou-se a execução da atividade de análise de assunto em três livros pré-selecionados com foco na atribuição de termos. A seleção dos materiais considerou apenas que os livros não pertencessem ao acervo da Universidade e que os temas fossem genéricos, não tendo sido consideradas na seleção as áreas de conhecimento. Na execução desta atividade foi incorporado o preenchimento de uma ficha baseada em Coutinho e Araujo (2010) com vistas a identificar os procedimentos realizados durante a tarefa.

Após a execução da atividade foi utilizada a técnica Protocolo Verbal, que consiste na verbalização consciente dos pensamentos dos indivíduos diante da execução de uma atividade com vistas a fornecer informações de seus processos mentais a partir da exteriorização por meio da fala. Por fim, foi solicitada a elaboração de um desenho baseado nos protocolos do AT-9.

3.1 Resultados

Na análise da realização da tarefa proposta verificou-se que a atuação dos bibliotecários, denominados S1, S2 e S3, não foi uniforme (QUADRO 1) variando os procedimentos executados conforme o livro analisado. Apenas a leitura do título e subtítulo foi o procedimento realizado por todos os bibliotecários em todos os livros analisados (100% de realização).

QUADRO 1 - Procedimentos realizados para definir a escolha dos termos

Procedimento	Livro 1			Livro 2			Livro 3		
	S1	S2	S3	S1	S2	S3	S1	S2	S3
Leitura do título e subtítulo	X	X	X	X	X	X	X	X	X
Nome do autor	X			X		X	X		X
Lombada		X							
Área do livro	X		X	X		X	X		X
Leitura do resumo		X			X				X
Leitura do índice		X							
Leitura do sumário	X		X	X		X	X		X
Leitura da introdução	X	X		X	X		X		
Leitura dos capítulos					X				
Leitura da orelha do livro	X		X	X	X	X			
Leitura dinâmica				X	X				

Folheada geral		X	X	X	X		X	X	
Ilustrações, diagramas, tabelas e seus títulos explicativos							X		
Leitura de palavras ou grupos de palavras em destaque (sublinhadas, impressas em tipos diferentes etc)	X								
Exame das referências bibliográficas	X						X		
Material adicional									
Catálogo na fonte	X			X		X	X		X
Consulta outro bibliotecário	X			X			X		
Outras (inseridas pelos bibliotecários)									
Consulta a outras bases de dados	X			X			X		
Rede Pergamum			X						
Catálogo da Universidade que publicou									X

Fonte: Dados de pesquisa. Elaborado pela autora

Interessante observar que, além das opções apresentadas no formulário, poucos foram os procedimentos adicionais realizados pelos bibliotecários para auxiliar na definição dos termos. Não foi mencionada pelos pesquisados a execução de práticas de apoio em contextos colaborativos, mas, no quadro acima, é possível verificar a utilização de outras bases de dados por um dos entrevistados e a consulta a um ambiente “coletivo” por parte de outro bibliotecário. Entretanto, essas práticas demonstram, a nosso ver, um aporte tímido aos ambientes construídos consensualmente por comunidades que visam o compartilhamento de significados uniformes numa abordagem difundida pela web semântica.

A outra perspectiva adotada na pesquisa, referente à análise dos termos atribuídos pelos bibliotecários na tarefa executada, revelou também resultados diferenciados:

- No primeiro livro houve certa homogeneidade na definição dos termos com a adoção das expressões “extensão universitária” e “ensino superior” pelos três bibliotecários;
- No livro dois a seleção de termos girou em torno de um eixo comum – Mudanças climáticas – apesar dos termos escolhidos pelos pesquisados terem sido variados como: Brasil - Clima (S1); Tempo (S2); Climatologia (S3);
- No terceiro livro os termos foram bem diversificados, não oferecendo um ponto comum que pudesse ser usado na recuperação deste livro em uma hipotética pesquisa realizada por um usuário. Foram apresentados termos como: Informática – estudo e ensino (S1), Programação de computadores – computação (S2) e Armazenamento de dados (S3).

3.2 Análise simbólica

A utilização da TIC possibilitou a observação, por meio do relato dos pesquisados, de situações críticas vivenciadas na atividade de análise de assunto. A vantagem do uso desta

técnica foi possibilitar a identificação de eventos relatados na perspectiva dos entrevistados de modo a compor a análise dos aspectos subjetivos numa perspectiva holística em que o indivíduo não é considerado isolado de seu contexto e história, o que permitiu nesta pesquisa um entrelaçamento entre fatos presentes e passados possibilitando um olhar profundo sobre o fenômeno estudado.

A partir das inferências sobre os incidentes relatados foi possível identificar os seguintes pontos críticos na situação apresentada por cada entrevistado (QUADRO 2):

QUADRO 2 - Incidente crítico – aspectos centrais

	Ponto central	Aspecto crítico
S1	Dúvida quanto ao termo a ser usado. Não existia o termo autorizado	O usuário não vai achar o livro
S2	Divergência com o usuário	Conflito de autoridade
S3	Diversidade de definições de um termo	Um termo pode ter vários significados dependendo do contexto

Fonte: Dados de pesquisa. Elaborado pela autora

Sobre os aspectos simbólicos, a consolidação das imagens características do exercício da catalogação e representativas do incidente crítico constam dos QUADROS 3 e 4.

QUADRO 3 - Imagens representativas da atividade de catalogação

	S1	S2	S3
Imagem	<i>“Livro aberto”</i>	<i>“Árvore”</i>	<i>“eu pegaria... o globo e o colocaria em movimento assim”</i>
Explicação	<i>“Sempre que eu penso em uma coisa ... boa, eu penso em um livro aberto..., sempre.”</i>	<i>“Porque a árvore está crescendo, dá frutos, dá folha; eu vejo como uma árvore mesmo, uma coisa produtiva.”</i>	<i>“... porque é movimento, é movimento ... de ideias, de assunto, de saberes, de...então... movimento, uma coisa que representasse o movimento, um globo”</i>

Fonte: Dados de pesquisa. Elaborado pela autora

QUADRO 4 - Imagens representativas do Incidente crítico

	Imagem	Explicação
S1	<i>“Ah, barreira, eu acho que um muro. rrsrrs, É, uma barreira, assim bem, sabe...”</i>	<i>“Ah, porque é, porque é uma barreira mesmo, sabe, eu acho que é, é a dificuldade que o usuário tem de localizar e a gente tem de disponibilizar desta forma.”</i>
S2	<i>“um menino correndo, correndo e você não conseguindo pegar ele ... um menino bem levadinho, aquele que você dá uns tapinhas nele e ele..., sai correndo desesperado.”</i>	<i>“Aí ... mais ou menos isso, porque nós não entramos em... num diálogo né?”</i>
S3	<i>“hum... imagem que vem para mim é muito de... antena,”</i>	<i>“A antena com essas... sempre... nunca sozinha... sempre com alguma coisa junto...”</i>

Fonte: Dados de pesquisa. Elaborado pela autora

Investigando-se com mais vagar as imagens propostas pelos entrevistados para representar a atividade de catalogação e os incidentes críticos pode-se observar, conforme as considerações apresentadas por Jean Chevalier, que os símbolos constituem o cerne da vida imaginativa, revelando segredos do inconsciente e traduzindo o esforço dos entrevistados em “decifrar e subjugar um destino que lhe escapa através das obscuridades que o rodeiam” (CHEVALIER e GHEERBRANT, 2008, p. XII). Estas percepções podem ser observadas nas significações que se “corporificaram” ao analisar o comportamento de cada entrevistado apresentadas a seguir:

Em S1 a atividade de catalogação é representada como um “livro aberto”. Cirlot (1984) trata do simbolismo do livro dizendo que os chineses consideram emblemas com livros como símbolos do poder para afastar espíritos malignos; já Chevalier e Gheerbrant (2008) destacam que os livros silibinos eram consultados pelos romanos em situações excepcionais para encontrar neles as respostas divinas para suas angústias. Os mesmos autores destacam que o livro dos mortos egípcio era utilizado para implorar aos deuses uma boa travessia dos infernos e a conseqüente chegada ao sol eterno. Interessantemente, ao representar o seu incidente crítico através da imagem de uma barreira a ser transposta, S1 cria, sem perceber, uma metáfora que se constela com o simbolismo do livro. Uma vez que a barreira (ou o muro) é descrita por Cirlot (1984) como uma incapacidade de avançar pelo caminho pretendido, o livro, do modo como foi citado (“*uma coisa boa*”) parece evocar um salvo-conduto ou proteção para afastar os “espíritos malignos” ou como fonte de “respostas divinas para as suas angústias” e prescrição para implorar uma “passagem pelos infernos” e o acesso ao “sol eterno”. Desse modo, parecem ficar bastante claras as associações, não somente do livro como um lenitivo para a angústia diante do desconhecido e do muro com uma barreira – um elemento que tem em sua significação mais fundamental o sentido de separação – como também entre essas duas imagens (livro e barreira/muro) na forma de Nêmesis um do outro. Isso revela muito da visão de S1 sobre o processo de catalogação e seus percalços. Para S1, parece tratar-se de uma atividade tranquila, que, porém, pode tornar-se turbulenta quando a impossibilidade de acesso se configura como uma separação, uma “comunicação cortada”, frustrando que o objetivo maior do entrevistado – que é disponibilizar para o usuário o acesso a informação da melhor forma possível – se concretize conforme seu intento.

Já para S2, a catalogação é representada como uma árvore, cujas interpretações principais de seu simbolismo articulam-se em torno da ideia de um cosmo vivo em perpétua regeneração, conforme mencionam Chevalier e Gheerbrant (2008). Como vida inesgotável, equivale à imortalidade e esse conceito de vida sem morte se traduz em realidade absoluta,

tornando-se a árvore como centro do mundo (CIRLOT, 1984). Tida como um dos temas simbólicos mais ricos e difundidos, a árvore consolida o “aspecto cíclico da evolução cósmica: morte e regeneração”, significação que é referenciada por Mircea Eliade assim retratada por Chevalier e Gheerbrant, (2008, p.85): “é por ser vertical, é porque cresce, pende suas folhas e torna a recuperá-las e porque, conseqüentemente, se regenera: morre e renasce inúmeras vezes.”

Este sentido encontra repercussão no imaginário de S2 que justifica a sua escolha por esse símbolo ao percebê-lo como algo que “*está crescendo, dá fruto, dá folha [...] uma coisa produtiva*”. Essa imagem de morte e regeneração relacionada à catalogação relembra seus processos em que o texto analisado “morre”, se “regenerando” nos termos atribuídos para representá-lo no final do processo de análise de assunto, numa atividade que abriga o sentido de produção e transformação. Também o verticalismo da árvore, que tem suas raízes fincadas no solo e seus galhos elevados para o céu, evoca-a como símbolo das relações entre o céu e a terra, interpretação vista nos mulçumanos xiitas de rito ismaelita, no qual a árvore simboliza *hakikal*, um estado de beatitude no qual o mítico, ao ultrapassar a dualidade das aparências, encontra a realidade suprema (CHEVALIER e GHEERBRANT, 2008).

Uma postura de “supremacia” pode ser inferida no posicionamento que S2 adota em relação ao resultado de seu trabalho, encarnado como uma “verdade absoluta”, que é refletido no sentido da árvore como “centro do mundo”, o que pressupõe que qualquer questionamento ou interferência externa não será visto com naturalidade. Essa constatação aparece explicitada na justificativa da evocação da imagem do incidente crítico (“*porque nós não entramos em... num diálogo né*”) na qual o questionamento a uma posição adotada por S2 e ausência de concordância com seu ponto de vista são considerados como fator de impasse, pois se choca com o sentido de “eixo do mundo”, “ pilar central”, simbolizado pela árvore, que constitui o significado da atividade de catalogação para o entrevistado e em torno do qual circundam seus comportamentos.

O reforço a esta interpretação pode ser associado também na imagem representativa do incidente crítico (“*menino correndo*”) no qual o confronto de opiniões – que gerou um conflito de autoridade e um questionamento da “verdade suprema”, que é a decisão tomada – configura-se como um ato de rebeldia e não aceitação de uma realidade adversa. Segundo Cirlot (1984, p.378), a imagem do menino representa “o produto da *coniunctio* entre inconsciente e o consciente”, definida como um símbolo alquímico de união de substâncias desiguais. A corrida remete ao sentido de fuga, o que pode caracterizar a resistência em proceder a união das “ideias opostas”, desiguais, fruto de pressão exercida pelo contexto, ao

qual S2 demonstra resistência e do qual gostaria de fugir. A imagem do menino, nas colocações de Chevalier e Gheerbrant (2008, p.302), remete à “vitória sobre a complexidade e a ansiedade e a conquista da paz interior e da autoconfiança” e sua fuga, na representação do incidente crítico, busca preservar esse estado de completude que está sendo ameaçado.

O terceiro entrevistado (S3), ao representar a atividade de catalogação, opta pela imagem de um globo em movimento. Este símbolo evoca em seu significado o sentido de poder, “o domínio ou o território sobre o qual se estende a autoridade do soberano e o caráter *totalitário* dessa autoridade” (CHEVALIER e GHEERBRANT, 2008, p.472). Outra associação importante desta imagem relaciona-se a sua forma esférica que, segundo Cirlot (1984), corresponde à perfeição e felicidade, pois a ausência de arestas (quinas) remete à falta de inconvenientes, estorvos e contrariedades. Baseado também neste autor articula-se ao símbolo a ideia de esfericidade relacionada a forma circular do globo que, inspirado nos emblemas chineses, representa atividade, símbolo do princípio masculino (*Yang*) representado por um círculo branco.

Como justificativa para a escolha da imagem, S3 associa o conceito de movimento (“*de ideias, de assunto, de saberes*”...), termo que reforça o simbolismo da perfeição como citado por Chevalier e Gheerbrant (2008, p.389): “A noção de esfera e de movimento orbicular é dominante e exprime a perfeição. Se um ser for concebido como perfeito, ele será simbolicamente imaginado como uma esfera”.

A representação da atividade de classificação como um globo em movimento sugere que S3 vê seu trabalho como uma atividade “perfeita”, sem arestas, o que possibilita compreender como a antena – imagem associada ao incidente crítico – pode ser responsável por desarticular esse mundo considerado pleno. A simbolização da antena neste contexto está associada ao inconstante, ao múltiplo, ao incontrolável em si mesmo, visto que pode absorver vários significados dependendo do cenário. Essa faceta que a antena representou constitui uma ameaça ao universo perfeito e controlado de S3 porque, apesar deste representar movimento, esse movimento é orbital e gira em torno de si mesmo e a dinamicidade representada pela antena não é linear, trazendo em sua representação o sentido do “imperfeito” e imprevisível.

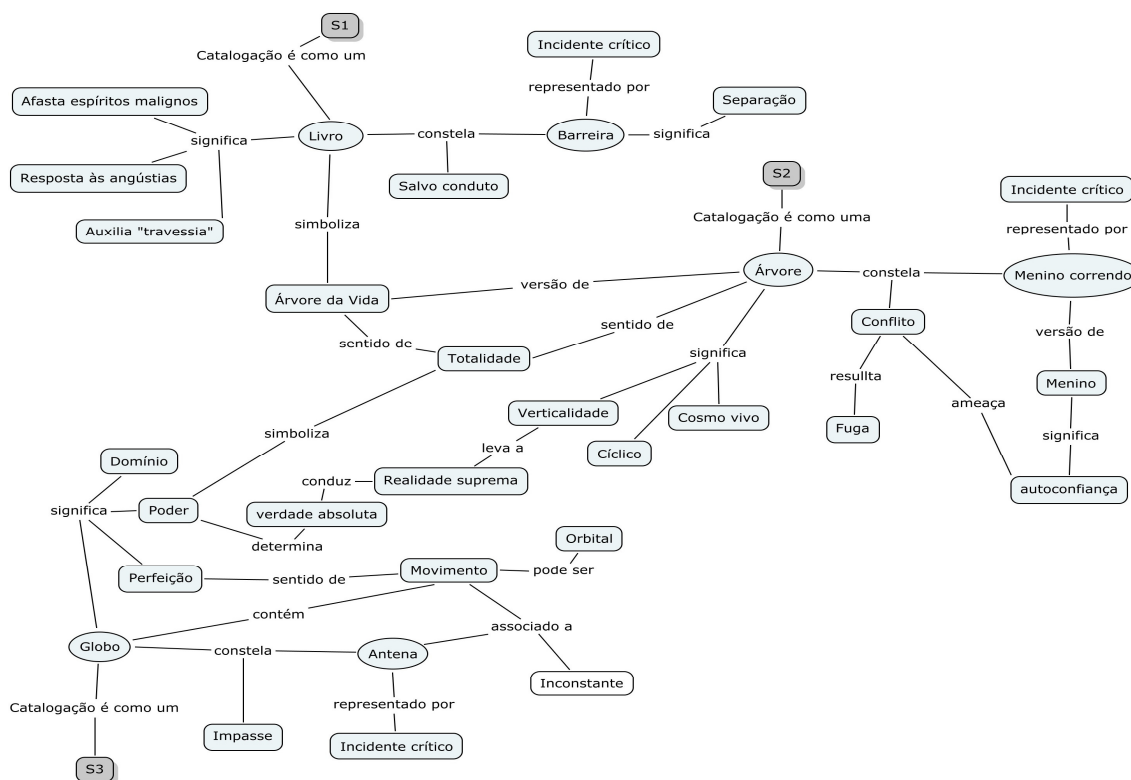
Ao reunir as associações feitas por S1, S2 e S3 pode-se perceber uma conexão simbólica estreita entre essas metáforas. Quando S2 e S3 evocam como representação da atividade de catalogação, respectivamente, uma árvore e o globo, verifica-se como o conceito de plenitude – traduzido na “verdade absoluta” e na “perfeição” – encontra repercussão no imaginário de ambos entrevistados, sugerindo que o produto do trabalho realizado é completo

em si mesmo. A mesma similaridade é vista nas imagens evocadas nos incidentes críticos caracterizadas pela quebra dessa hegemonia expressa tanto na figura de outra opinião, quanto na dinamicidade de significados que um termo pode assumir. As duas situações representam um conflito que ameaça esse conceito de “supremacia” que permeia a visão dos entrevistados sobre a realização de suas atividades. O que ameaça a ambos é algo que conteste ou abale suas certezas.

Interessante associar que o livro, metáfora atribuída à catalogação por S1, tem, dentre seus significados, o relacionado ao “Livro da Vida do Apocalipse” onde este se identifica com a Árvore da Vida sendo “as folhas da árvore como os caracteres do livro” representando “a totalidade dos decretos divinos” (CHEVALIER e GHEERBRANT, 2008, p.555).

As associações entre os diversos sentidos atribuídos aos símbolos permitiram construir uma rede de significados (FIG. 1) cuja sutura se dá pela interpretação da catalogação como uma atividade que tem em si um aspecto de finitude, sendo seu resultado tido como um “produto fechado”, total e completo que, apesar de ser construído sob os olhares e caminhos diferentes por cada um dos entrevistados, para cada um deles o resultado de seus trabalhos adquire um caráter de verdade “incontestável”.

FIGURA 1 - Rede interconectada de símbolos e significados



Fonte: Elaborado pela autora

Segundo Cirlot (1984), a doutrina de *Mohyiddin ibn Arabi* propõe o universo como uma condensação do imenso livro escrito pela pena divina e que o universo transcendente do livro descende para tomar a forma do universo manifesto onde o homem vive. Esse simbolismo, destaca o autor, é o mesmo que os Rosacruztes chamavam de *Liber Mundi* e o Apocalipse de *Liber Vitae*. Já Chevalier e Gheerbrant (2008) fazem referência ao livro (*Liber Mundi*) não somente como uma simbolização da relação/correspondência, via mensagem divina, do macrocosmo com o microcosmo, como também faz uma referência a versões da busca do Graal onde o livro seria uma representação do próprio Graal cuja busca seria a “procura da palavra perdida”, da sabedoria tornada acessível ao comum dos mortais. Neste caso, a conexão simbólica com a atividade de catalogação parece inquestionável. Os autores destacam ainda que enquanto o livro fechado representa a matéria virgem, o “livro aberto” representa a matéria fecundada. Ou seja, enquanto o livro permanece fechado, o mesmo conserva seus segredos, já quando é aberto seu conteúdo é tomado por quem o investiga. Não seria essa a essência da atividade de catalogação? Tornar abertos os livros para permitir que os investigadores tenham acesso às suas mensagens? Ainda fazendo referência aos autores retro mencionados, cabe à catalogação repetir o trabalho dos alquimistas que, através de suas operações sucessivas, abrem os selos e as fitas que fecham o *Grande Livro da Natureza* e permitem que a matéria prima seja extraída da mina e trabalhada. Para que isso ocorra é necessário que, dos conteúdos intrínsecos dos livros, seja retirada sua essência de modo que possam ser representados fielmente.

Esta atividade encontra analogia no sentido de verticalidade da árvore, apresentado por Chevalier e Gheerbrant (2008), que coloca em comunicação os três níveis do cosmo: o subterrâneo, a superfície e as alturas. No processo de catalogação, esses níveis podem ser refletidos, respectivamente, no texto do livro, nos termos atribuídos e no usuário da informação: por meio das raízes se acessa as profundezas onde se enterram as essências; na superfície da terra o tronco, representação visível da estrutura, vai levar aos galhos e folhas a seiva (informação) necessária para saciar e atender a necessidade daqueles que estão no final do processo. “Ela estabelece, assim, uma relação entre o mundo ctoniano e o mundo uraniano” e reúne todos os elementos (CHEVALIER e GHEERBRANT, 2008, p.84).

A finalização desse ciclo completo remete ao significado contido no globo que, por sua forma esferoidal, corresponde à plenitude da realização. Seu aspecto circular é associado às propriedades simbólicas da perfeição e homogeneidade, atributos que se espera que a atividade de catalogação compreenda que é a perfeita representação do livro e uma codificação que traduza de maneira homogênea e fiel seu conteúdo. Na tradição grega, em

especial em Parmênides e nos textos órficos, as esferas são utilizadas para representar dois mundos: o mundo terrestre e o Outro-Mundo, sendo a morte responsável por fazer passar de uma esfera a outra. Também nesse sentido a catalogação pode ser vista como representativa da passagem de um mundo do conteúdo para um mundo de conceitos, no qual o primeiro é “abandonado/eliminado” após a inserção no novo mundo que passará a representar o mundo desejável.

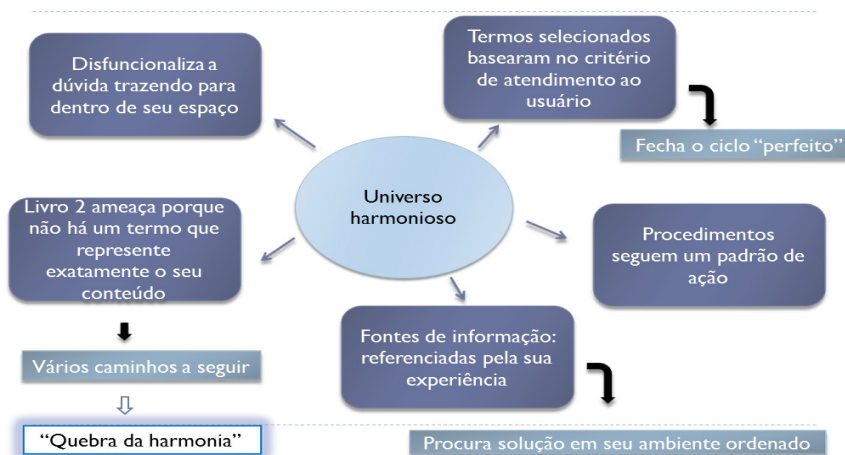
É importante destacar que essa análise não se esgota na mera interpretação dos significados latentes, pois as palavras, segundo Jean Chevalier, são incapazes de expressar todo o valor de um símbolo (CHEVALIER e GHEERBRANT, 2008, p. XIII). Cabe nessas reflexões o cuidado de não encerrar em estreitos limites todas as dimensões descortinadas pela interpretação simbólica, mas ampliá-la em outras metodologias que possam expandir seus significados, proposta que foi complementada nesta pesquisa com o uso do AT-9 por meio do qual se buscou compreender o que permeia as ações dos sujeitos em situações de angústia (CARDOSO, 2005) representada neste estudo pela decisão do termo adequado na análise de assunto.

O uso do Teste Arquetípico de Nove Elementos (AT-9) e de outras técnicas complementares, como o incidente crítico e o uso da expressão criativa possibilitou explorar os aspectos subjetivos utilizando-se as dimensões simbólico-afetivas. A compilação dos dados obtidos foi realizada sob o crivo das estruturas antropológicas do imaginário de Gilbert Durand (1997), consolidada na metodologia desenvolvida por Yves Durand (1988) com vistas a identificar indícios que revelassem como se deu o processo informacional, subjetivo e afetivo envolvido na atividade executada.

Como resultado desta etapa foi possível identificar os microuniversos míticos de cada indivíduo e suas relações com os comportamentos informacionais adotados, o que possibilitou compreender o porquê dos resultados diferenciados na execução da atividade de análise de assunto.

O perfil místico de S1 o caracteriza como um indivíduo que procura disfuncionalizar a ameaça trazendo-a para seu universo controlado e submetendo-a a sua experiência de forma a neutralizar a angústia que ela encerra. Esse perfil direciona suas atitudes no processo de tratamento informacional e os termos selecionados na atividade de análise de assunto fecham um ciclo “perfeito”, no entender do pesquisado, que é atender ao usuário (FIG. 2).

FIGURA 2 - Microuniverso de S1 e ações decorrentes



A identificação do perfil de S2 como sintético simbólico, no qual o movimento cíclico é traduzido pelo “eterno retorno”, permitiu vislumbrar que o que caracteriza o pesquisado é sua postura fechada para decidir. Esta atitude torna seu processo decisório bloqueado a interferências externas, sendo esse modo de se proteger em si mesmo sua forma de lidar com os desafios da decisão (FIG. 3). Os termos selecionados na atividade de análise de assunto seguiram o perfil de S2 de procurar em um ambiente interno – ou seja, no próprio livro – a resposta à atividade, mantendo nessa postura o ambiente hermético.

FIGURA 3 - Microuniverso de S2 e ações decorrentes



Em relação a S3 foi possível inferir que seu perfil sintético existencial diacrônico o caracteriza como um indivíduo que busca lidar com a decisão sob duas perspectivas: volta-se para si para encontrar respostas e, em seguida, parte para uma postura de enfrentamento utilizando os instrumentos de que dispõe para enfrentar a dúvida e decidir (FIG.4). Na atribuição de termos, o pesquisado trouxe para si o processo de análise, mas, na hora de

decidir, utilizou um instrumento externo capaz de referendar de forma pragmática sua decisão.

FIGURA 4 - Microuniverso de S3 e ações decorrentes



4 CONCLUSÃO

As técnicas e métodos adotados permitiram concluir que os aspectos subjetivos – mesmo em um processo que tem uma metodologia estruturada e formalizada em termos de vocabulários controlados, normatizações de procedimentos e fontes de informação padronizadas como é o caso da análise de assunto – são responsáveis por resultados diferenciados. Isto ocorre porque o processo de seleção de termos, que é uma atividade de tomada de decisão, não carrega em si apenas aspectos racionais, mas é perpassada por toda uma história de vida, experiências profissionais, preferências, estruturações mentais e perfis psicológicos que fazem desse processo o resultado de toda uma confluência de fatores que não se situam apenas em nível consciente, estando influenciado por circunstâncias que não estão palpáveis ou explícitas no comportamento do indivíduo.

As constatações possibilitadas pela pesquisa permitiram compreender a complexidade do processo de tratamento da informação e caminhar na busca pelo estabelecimento de padrões e critérios mais claros e diretrizes mais precisas de modo a possibilitar o estabelecimento de parâmetros mais eficazes para orientar o processo de indexação.

Acredita-se que o reflexo de um processo de tratamento informacional balizado em critérios mais efetivos e que observem a subjetividade como fator intrínseco à atividade irá culminar em uma recuperação de informação mais precisa contribuindo para que o usuário tenha atendida sua demanda de forma satisfatória.

Conforme destacado por Ferneda (2003) a recuperação da informação implica em operar seletivamente um estoque de informação, o que envolve processos cognitivos por parte dos usuários, bem como a capacidade de abstração, apreensão e representação do significado de forma contextual por parte dos indexadores. Essas particularidades dificilmente podem ser formalizadas por meio de um algoritmo, mas devem ser observadas sob outras perspectivas visando resultados mais efetivos nesse processo.

REFERÊNCIAS

- ARAÚJO, E. P. O. **Tomada de decisão organizacional e subjetividade**: análise das dimensões simbólico-afetivas no uso da informação em processos decisórios. 2013. 162 p. Dissertação (Mestrado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais. Belo Horizonte. 2013.
- BORKO, H. Information Science: What is it? **American Documentation**, v.19, n.1, p. 3-5, 1968.
- CEMIN, A. B. et al. Gênero e imaginário. **Revista eletrônica do Centro de Estudos do Imaginário**. Ano I, n.3, out-dez. 2001. Disponível em <http://www.cei.unir.br/artigo32.html>. Acesso em 18 mai 2014.
- CHEVALIER, J.; GHEERBRANT, A. **Dicionário de símbolos**. 22. ed. Rio de Janeiro: José Olympio, 2008
- CIRLOT, J. E. **Dicionário de símbolos**. São Paulo: Editora Moraes, 1984.
- COUTINHO, L. F.; ARAÚJO, C. A. A. A indexação nas áreas do conhecimento: uma comparação das áreas de ciências exatas e da terra, das ciências humanas e da linguística, letras e artes. Rio de Janeiro. 2010 **Anais**. Rio de Janeiro. XI ENANCIB. 2010
- DURAND G. **As estruturas antropológicas do imaginário**. Martins Fontes. São Paulo. 1997.
- DURAND, Y. L'exploration de L'imaginaire: Introduction à la modélisation des Univers Mythiques. L'espace bleu. Paris. 1988.
- ESTRADA, A. A.. O teste AT-9 na escola: considerações preliminares acerca do universo da angústia. **Educere. Revista da Educação**. v.2, n.1, p. 25-38, jan/jun 2002.
- FLANAGAN, J. C. A técnica do incidente crítico. **Arquivos brasileiros de Psicologia Aplicada**, v.25, n.2, abr/jun 1973.
- FERNEDA, E. **Recuperação de informação**: análise sobre a contribuição da Ciência da Computação para a Ciência da Informação. Tese (Doutorado em Ciências da Comunicação). Escola de Comunicação e Artes. Universidade de São Paulo. São Paulo. 2003.
- FUJITA, M. S. L. 2003. **A leitura documentária do indexador**: aspectos cognitivos e linguísticos influentes na formação do leitor profissional. Tese (Livre-docência em Análise Documentária e Linguagens Documentárias Alfabéticas). Universidade Estadual Paulista. Marília. 2003.

- FUJITA, M. S. L et al. 2010. O contexto da indexação para a catalogação de livros em abordagem sociocognitiva. **Brazilian Journal of Information Science**, Marília (SP), v. 4, n. 2, p. 22-40, jul/dez 2010. Disponível em:
<http://www2.marilia.unesp.br/revista/index.php/bjis/index>. Acesso em: 8 mai 2014
- GARCIA, J. C. R, Recuperação da informação. In: **DataGramZero** - Revista de Ciência da Informação, v. 8, n. 6, 2007. Disponível em http://www.dgz.org.br/dez07/F_I_com.htm. Acesso em 20 mai 2014
- HARRIS, J. L. **Subject analysis**: computer implications of rigorous definition. Metuchen, NJ: Scarecrow Press, 1970
- How much information 2009?** Disponível em:
http://hmi.ucsd.edu/pdf/HMI_2009_ConsumerReport_Dec9_2009.pdf. Acesso em 15 mai 2014
- KOBASHI, N. Y. **Elaboração de informações documentárias**: em busca de uma metodologia. Tese (Doutorado em Ciências da Comunicação). Escola de Comunicações e Arte, Universidade de São Paulo. São Paulo. 1994.
- LANCASTER, F.W. **Indexação e resumos**: teoria e prática. 2 ed. Briquet de Lemos. Brasília. 2004
- NAVES, M. M. L. Análise de assunto: concepções. **Revista Biblioteconomia**. Brasília, v.20, n.2, p.215-226, jul/dez. 1996.
- OLIVEIRA, T. M. V. Amostragem não probabilística: adequação de situações para uso e limitações de amostra por conveniência, julgamento e quotas. **Administração online**, v.2, n.3, jul/ago/set 2001. Disponível em http://www.fecap.br/adm_online/art23/tania2.htm
- OLIVEIRA, G. P.; MAIA, L. S. L. **Estudo do universo imaginário de professores de matemática**: uma análise a partir da teoria de Gilbert Durand. 2008. Disponível em:
<http://31reuniao.anped.org.br/1trabalho/GT19-4798--Int.pdf>. Acesso em: 20 mai. 2014.
- PAULA, C. P. A. Dimensões simbólicas e afetivas do uso da informação: uma análise das comunicações entre professores do departamento de psicologia de uma instituição de ensino superior pública brasileira. Brasília. 2011. **Anais**. Brasília. XII ENANCIB. 2011.
- PAULA, C. P. A. Proposta de metodologia para a investigação do comportamento de busca informacional e do processo de tomada de decisão dos líderes nas organizações: introduzindo uma abordagem clínica na informação. Rio de Janeiro. 2012. **Anais**. Rio de Janeiro. XIII ENANCIB. 2012
- SILVA, M.R. et al. 2004. A prática de indexação: análise da evolução de tendências teóricas e metodológicas. **Transinformação**, v.16, n.2, p-133-161, mai/ago 2004. Disponível em <http://revistas.puc-campinas.edu.br/transinfo/viewarticle.php?id=65>. Acesso em 09 mai 2014
- RUBI, M. P.; FUJITA, M. S. L. Política de indexação na catalogação de assunto em bibliotecas universitárias: a visão sociocognitiva da atuação profissional com protocolo verbal. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v.7, n. 2, p. 118-150, jan./jun. 2010
- UNISIST. Princípios de indexação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v.10, n.1, p. 83-94, mar. 1981.

ORGANIZAÇÃO DA INFORMAÇÃO JURÍDICA: O ARQUIVO DA JUSTIÇA FEDERAL DA PARAÍBA (JFPB), USOS E USUÁRIOS.

ORGANIZATION OF LEGAL INFORMATION: ARCHIVE OF FEDERAL COURT OF PARAIBA (JFPB) USE AND USERS

Wendia Oliveira de Andrade
Dulce Amelia de Britto Neves

Resumo: Nas unidades informacionais é comum a presença de usuários que fazem uso dos mais diversos tipos de serviços e/ou produtos por elas oferecidos. O Arquivo da Justiça Federal da Paraíba apresentou-nos usuários passíveis de investigação e tipificação, colaborando para que os estudos de usuários ganhem cada vez mais espaço no contexto arquivístico. Objetivamos conhecer esses usuários e descrever os tipos existentes, suas necessidades e descrever suas estratégias de busca pela informação contida nos documentos que encontram-se no Arquivo Judicial. Para isso, fizemos uso de métodos que fossem além de números, pois nossa pesquisa caracteriza-se pelo uso da Fenomenologia, como metodologia para obter, descrever e analisar as informações obtidas através da observação, entrevistas guiadas e do Protocolo Verbal. A ênfase da pesquisa foi dada ao usuário, mas chamou-nos atenção nos discursos dos usuários que a forma de representar e organizar a informação tornou-se muitas vezes elementos de barreira ao acesso e uso das informações processuais, tanto físicas quanto no SRI *Tebas*. A pesquisa possibilitou essas conjecturas sobre acesso, uso, usuário, representação da informação devido a união interdisciplinar entre a Ciência da Informação (recuperação da informação), a Arquivologia (campo onde desenvolveu-se a pesquisa, e continha elementos específicos dos arquivos) e da Biblioteconomia (estudos de usuários). Assim, destacamos nesse estudo não apenas os usuários da JFPB, mas também a necessidade de representar melhor a informação, com intuito de recuperá-la com eficiência e eficácia.

Palavras-chave: Arquivo Judicial. Usuário de Arquivo. Organização da Informação Jurídica. Fenomenologia, Protocolo Verbal

Abstract: In the informational units is common to have users who make use of several types of services and / or products offered by them. In the specific files case, the presence of users is not always constant to be easily observed. In the Judicial Archive of the Federal Court of Paraíba we can note this difference, the existence of users in bigger quantity than most legal files of the capital of Paraíba , and it is for this reason that we are in a search of exploring , investigate and meet those who seek information in the Judicial Archives environment. We aim to meet these users and describe the types that exist, their needs and search strategies for the information contained in the documents that are in the Archive Judiciary. Therefore it was used methods that were beyond numbers, because our research is characterized by the use of phenomenology as a methodology to obtain, describe and analyze the information through interviews with the subjects of the research and Verbal Protocol , used to describe the face to face interaction of the user and professional /manager of Archive information. The use of the Information Science of the Librarianship and Archival served as the basis for all the informational construct of the meanings and significance of the research. Emphasis was given to the user, but also the key role of information for him/her, as active subjects to access and use this process.

Keywords: Judicial Archive, Organization of legal information, Archive User. Phenomenology, Verbal Protocol Analysis.

1 INTRODUÇÃO

Pensar em um conceito de informação e em todos os aspectos que o configuram e o caracterizam num determinado ambiente informacional, nos parece algo bastante complexo. Seja pelas inúmeras definições encontradas para o termo, sintática e semanticamente falando, seja por toda a carga de significações que agrega, ou simplesmente pelo fato de a empregarmos em determinado contexto.

Mas a informação de que tratamos especificamente neste trabalho possui características, aspectos e objetivos a serem observados, estudados e descritos, pois a informação que observamos é aquela que **se evidencia** no momento da busca informacional, em um Arquivo Público Federal, de documentação específica e de informações precisas, legais e comprobatórias. Passa a ser então esta informação mais importante do que as demais, ou seja, aquelas informações que aparecem em outras unidades informacionais? Ou em outros contextos de utilização e acesso? A resposta certamente é negativa, tendo em vista que toda informação possui sua relevância no ambiente que permeia. Observando o Arquivo Judicial, enquanto se desenvolvia um estágio extracurricular no curso de graduação em Biblioteconomia, percebemos seu potencial informacional. Chamou-nos a atenção ser um Arquivo Jurídico, com documentação específica, com uma organização informacional, representação, linguagem e tratamento documental próprios e com uma busca informacional tão intensa - tanto por parte dos próprios servidores da instituição, quanto dos jurisdicionados, ou seja, aquelas pessoas comuns que tinham algum processo sob a jurisdição da JFPB.

O arquivo, enquanto unidade informacional, muitas vezes possui um número reduzido de usuários, estes em sua maioria buscam-no para a realização de pesquisas, atividades como visitas dirigidas, entre outras; entretanto o Arquivo Judicial destacava-se por ser objeto de uma grande procura por parte de seus usuários e os próprios servidores da JFPB.

Só com um maior aprofundamento da Ciência da Informação (CI), constatamos que era este o campo (SARACEVIC, 1996) que iria nos dar todo o aporte para explorarmos os problemas que se apresentavam, possibilitando assim a efetivação de uma pesquisa com cunho informacional, mas que perpassasse outros campos do conhecimento, agregando valor e conteúdo à pesquisa.

A Arquivologia cedeu-nos de bom grado as especificidades da unidade informacional, seus elementos e instrumentos para organização e classificação (como a Teoria das três idades, a tabela de temporalidade (TTD) e sua lei de acesso a informação); a Biblioteconomia, com sua grande contribuição nos estudos de usuário e importância da indexação e cuidado ao representar a informação para posterior recuperação; e a CI com todo o seu arcabouço

informacional e todos os aspectos relevantes ao acesso, uso, recuperação e também nas questões de comportamento informacional para com o seu usuário.

Unimos estes grandes campos do conhecimento para buscar elucidar o problema de pesquisa que descrevemos em forma de questionamento: **com base nas necessidades apresentadas ao Arquivo Judicial por seus usuários, é possível identificá-los assim como as suas necessidades informacionais?** Agregamos em nível de reflexão à questão norteadora as formas de busca pelos usuários do Arquivo Judicial e sua **perspectiva do processo, enquanto documento e sua representação informacional** processual, ou seja, a busca realizada tinha quais elementos representativos.

O foco de pesquisa foi a identificação dos usuários existentes no Arquivo Judicial, o acesso e uso das informações obtidas por eles no ambiente informacional, entretanto enfatiza-se nesse trabalho a forma como os usuários buscavam a informação processual, ou seja, como a informação configurada num suporte impresso tinha sua representação compreendida por eles, e como esta se organiza dentro do próprio fluxo informacional do Arquivo.

Assim sendo apresentamos os objetivos da pesquisa, sendo o objetivo geral **identificar os usuários existentes no Arquivo da JFPB e as estratégias utilizadas na busca informacional** e como objetivos específicos, pontuamos aqueles que, tornando-se operacionalizáveis, nos seriam de grande ajuda: traçar o perfil dos usuários, com base na utilização dos serviços do Arquivo; verificar a existência de diferentes tipos de usuários; identificar suas necessidades informacionais (NI's); descrever as estratégias e o fluxo informacional dos usuários no Arquivo.

A metodologia de análise usada em todas as informações obtidas no campo da pesquisa está centrada na Fenomenologia, que descreve e estuda um fenômeno, essencialmente os fenômenos onde o homem é o centro de ação, no caso da nossa pesquisa: *o usuário em busca da informação*. Através da fenomenologia busca-se a interpretação do mundo ou do fenômeno através da percepção do sujeito (DARTIGUES, 2005), ou seja, considerando toda a sua carga informacional, toda a experiência adquirida. Segundo este método, todas as informações obtidas são importantes serem observadas, descritas e acima de tudo, compreendidas tal como se apresentam ao sujeito envolvido no processo.

Como técnica de coleta das informações, optamos pelo uso da entrevista e do protocolo verbal. Estas técnicas nos foram úteis para que pudéssemos analisar as informações passadas individualmente (entrevistas) e em interação face a face (protocolo verbal utilizado entre o profissional do Arquivo Judicial e o usuário, no momento da busca informacional).

2 A CIÊNCIA DA INFORMAÇÃO E A ARQUIVOLOGIA: REFLEXÕES BASILARES

2.1 Sobre a Ciência da Informação

Sobre a origem da CI, podemos elencar muitos fatores que culminaram com o seu surgimento e fundamentação enquanto ciência. Há relatos que apontam fatos históricos, como: a invenção da imprensa por Gutenberg; a explosão informacional no pós-guerra (segunda guerra mundial); problemas relacionados à organização informacional, a qual apresentava-se cada vez mais crescente e variada; como disponibilizá-la a tantas pessoas (a quem chamamos de usuários) que a quisessem para os mais diversos fins, considerando que a noção do termo usuário da informação passou a ser tratado tempos depois do surgimento da CI, quando o olhar social se fez presente e necessário a essa importante ciência.

O que não se pode ignorar é que o Tratado de Documentação (1934) de Paul Otlet e Henri La Fontaine trouxeram sob a nomenclatura de documento, o que consideramos atualmente como informação e todos os seus aspectos importantes abarcando desde a organização, armazenagem, recuperação até o uso da informação contida no interior intelectual dos *documentos*, esse último entendido aqui como o suporte. Não há dúvida de que foi grande a contribuição desses homens e cientistas da informação, que estavam à frente do seu tempo contribuindo para o surgimento de uma nova ciência – a CI – que contemplasse problemas de informação, independentemente do seu suporte. (FREIRE, 2006; ORTEGA; LARA, 2010).

Para tanto, afirma Garcia (2002, grifo nosso) sobre a gênese da CI, que vários autores:

[...] em publicações mais recentes ou mais antigas, dentre os quais **Pinheiro, 1998, Shera, 1968 e Taylor, 1966**, referem-se às conferências do *Georgia Tech* como marco para a ciência da informação. [...] Diversamente muitos autores citam **Borko (1968)**, adjetivando de clássica a definição que ele apresenta de ciência da informação. Porém no artigo em que conceitua ciência da informação ele afirma que faz uma síntese das definições de **Taylor (1966)**, publicada no *Annual Review of Information Science and Technology* (ARIST).

Sobre o objeto de estudo da CI, González Gómez (2000, p. 6, grifo nosso) afirma que:

Seja qual for a construção do objeto da Ciência da Informação, **ele deve dar conta do que as diferentes disciplinas, atividades e atores sociais constroem, significam e reconhecem como informação**, numa época em que essa noção ocupa um lugar preferencial em todas as atividades sociais, dado que compõe tanto a definição contemporânea da riqueza quanto a formulação das evidências culturais.

Com base em Araújo (2009), há pelo menos seis correntes teóricas que compuseram a CI, cada uma apresentando e agregando contribuições para a conceituação de informação e questionando aspectos a serem observados e tratados pela CI, elencam-se: os estudos de

natureza matemática (recuperação da informação e a bibliometria), a teoria sistêmica, a teoria crítica, as teorias da representação, os estudos em comunicação científica e os estudos de usuários.

É preciso compreender que, por mais que a CI tenha em seus fundamentos estruturais a recuperação da informação, o trato com profissionais e usuários da informação também é tema importante, não apenas como interação estes sujeitos com SRI's, mas como compreendem e buscam a informação. Consideramos então outra faceta importante da CI que é o tratamento de sua responsabilidade social. Como asseveram Freire e Freire (2010, p. 18):

Quando cientistas e profissionais da informação organizam textos ou documentos para atender a necessidade de um determinado setor da sociedade, o fazem acreditando que essas informações serão úteis para seus usuários potenciais e que, delas, resultarão benefícios para a sociedade.

Por isso, mais que uma CI que trata de problemas relacionados à informação, a CI busca consolidar seu caráter social, seja numa unidade informacional como o Arquivo Judicial, seja no auxílio aos profissionais e usuários da informação que ali se encontram buscando e fazendo uso da informação ali existente.

2.2 O registro das informações e a Arquivologia

A necessidade de registrar informações e acessá-las numa posteridade remete-nos aos primórdios das civilizações humanas, com certo grau de consciência da importância de tais registros. Todo o conhecimento que instigava e era descoberto, toda e qualquer informação que merecesse uma maior atenção era devidamente registrado (não importando quão rústico era o suporte a ser grafado ou inscrito) para que a consulta, o acesso e uso dos registros informacionais ali descritos pudessem ser vistos e revistos, ainda mais: ajudando a construir a memória de um povo. (LODOLINI, 1993).

Silva et al., (2009, p. 52, grifo do autor) nos remete a Hamurábi, para exemplificar a importância dos arquivos e mostrar os aspectos informacionais, já existentes naquela época:

Eles constituíam já um complexo **sistema de informação**. Para além, dos documentos em si, havia uma estrutura organizacional, um critério seletivo de preservação e a disponibilização de um serviço, determinado tanto pelo valor informativo das placas, como pela pertinência e rigor da sua integração sistêmica.

O local de guarda dessas informações era o Arquivo. Este local não era apenas o receptor e guardião de informações (em seus mais diversos suportes), mas o organizador desta memória. Nos arquivos eram guardados os mais importantes tesouros de uma sociedade: “os resultados de suas atividades políticas, sociais, econômicas, religiosas e até mesmo de suas vidas particulares.” (PAES, 2004, p. 15).

Todas essas informações distintas compunham uma sociedade, uma história, de povos que sentiam a necessidade de perpetuar sua memória para aqueles que viessem depois e assim constatassem a sua própria evolução, enquanto sujeitos cognoscentes, fizessem suas próprias análises e conjecturas a respeito daquelas informações ali depositadas.

Para tanto, não podemos negar que inicialmente os critérios utilizados para a guarda de tais materiais eram mais subjetivos, ou seja, guardar tudo o que fosse considerado importante sem maiores critérios de seleção, tornando-se um acervo com diversos tipos de suporte; mas com a chamada “explosão informacional” o fluxo informacional exigido e as características dos documentos passaram a exigir por si só, uma maior seletividade; provocando assim, um “aperfeiçoamento das técnicas de registro e análise dos documentos” (PAES, 2004) fazendo com que houvesse uma melhor adequação das atribuições administrativas e técnicas tanto para os arquivos, quanto para outras unidades informacionais que auxiliassem esse processo de organização, acesso e uso dos documentos e conseqüentemente das informações neles contidos. Cita-se como exemplo as bibliotecas, os centros de documentação/informação e os museus.

O arquivo, enquanto entidade mantenedora passou a ser então mais específico: em seus documentos, em sua forma de organização e em outras técnicas que o caracterizam singularmente como tal. Neste tocante, faz-se necessário apresentarmos uma definição de documento de arquivo:

Todos os livros, papéis, mapas, fotografias ou outras espécies documentárias, independentemente de sua apresentação física ou características, expedidos ou recebidos por qualquer entidade pública ou privada no exercício de seus encargos legais ou em função das suas atividades e preservados ou depositados para preservação por aquela entidade ou por seus legítimos sucessores como prova de suas funções, sua política, decisões, métodos, operações ou outras atividades, ou **em virtude do valor informativo dos dados neles contidos**. (SCHELLENBERG, 2006, p. 41, grifo nosso).

E é com base neste grifo do valor informativo dos documentos, em especial os que pertencem aos arquivos, apresentado por Schellenberg (2006) que nossa pesquisa começa sua relação direta com CI.

O potencial informacional apresentado nos documentos de arquivo, na questão de organização, acesso e uso, dentro e fora da instituição que os mantêm, apresenta-se como algo de extrema relevância de serem observados, estudados e descritos fundamentados pela CI, já que esta se encarrega dos mais diversos problemas, nas mais distintas áreas do conhecimento, desde que, estes partam da premissa primeira de ser questões do trato informacional. Como corrobora Freire (2006, p.6): [...] “o objeto de estudo da ciência da informação, a informação,

é um fenômeno que está relacionado a todos os campos do conhecimento científico, moldando-se aos interesses de cada uma delas.”

Então, com a CI percebemos o arquivo, não apenas como uma entidade de grande responsabilidade social para com seus documentos, servidores e possíveis usuários, mas como uma unidade possuidora de elementos informacionais com grande possibilidade de investigação e desenvolvimento de pesquisas e trabalhos extremamente ricos.

Sob a égide de unidade informacional, o arquivo e a Arquivologia, cedem à CI aspectos a serem observados com maior atenção:

[...] a Arquivologia trouxe para a Ciência da Informação não apenas um rico acúmulo de técnicas, conceitos e visões sobre os documentos arquivísticos, mas também ideias sobre a organicidade e o ciclo de vida dos documentos, sobre patrimônio e memória, sobre a historicidade dos registros do conhecimento humano para todas as disciplinas científicas pertencentes ao campo das ciências humanas e sociais [...]. (ARAÚJO, 2011, p. 119).

Observamos a importância da interdisciplinaridade desses campos do conhecimento trabalhando em consonância por um objetivo comum: explorar e aprofundar os aspectos informacionais congruentes apresentados tanto na Arquivologia como na CI. Os estudos destes campos interdisciplinares, identificam “a informação como elemento central do conjunto de objetos de que ambas se ocupam.” (FONSECA, 2005, p.10).

Com isso, consideramos as especificidades de cada campo do conhecimento, sem renegar de maneira alguma seu ponto comum: a informação. Neste sentido, utilizamos e adequamos conceitos de informação e de CI à nossa pesquisa, considerando que estes conceitos podem ser modificados para um melhor ajuste ao que se pretende pesquisar, sem com isso, descaracterizá-los.

3 A JUSTIÇA FEDERAL DA PARAÍBA: CAMPO DA PESQUISA

O Arquivo Judicial é público e federal⁷⁹, caracterizando-se como tal, devido às suas funções jurídico-administrativas para com o Estado de Direito, para com a população, e todos aqueles que figuram em seu ambiente informacional.

Podemos descrever como característica bastante peculiar do Arquivo Judicial, aquele que dentro da esfera federal, está ligado diretamente ao poder judiciário, o qual devemos compreender que sua função é de:

[...] interpretar e aplicar normas que compõem o **ordenamento jurídico nacional e distribuir a justiça entre as pessoas** (físicas e jurídicas). [...]

funciona como um árbitro da sociedade, julgando os conflitos de interesse com base nas constituições federais e estaduais, na Lei Orgânica dos Municípios, nas demais leis do país, nos costumes e em outros princípios do direito. (PASSOS; BARROS, 2009, p. 75, grifo nosso).

Passos e Barros (2009, p. 29) complementam que é o Estado “[...] quem deve exercer esse papel, e o faz por meio do poder judiciário. Todo indivíduo, sem distinção, tem o direito de reclamar a prestação jurisdicional.”

Se o Arquivo Judicial está subordinado a este tipo de poder, vislumbramos sua importância, tanto como parte de uma instituição séria e que funciona com bases legais, quanto por suas próprias diretrizes de funcionamento e tipo de documento que guarda.

Em conformidade com a Lei de Acesso à Informação (LAI) número 12.527 de 2011:

[...] precisamos lembrar que o poder público deve garantir uma gestão transparente da informação, propiciando amplo acesso a ela e sua divulgação; proteção da informação, garantindo-se sua disponibilidade, autenticidade e integridade; proteção da informação sigilosa e da informação pessoal, observada a sua disponibilidade, autenticidade, integridade e eventual restrição de acesso.

No caso do Arquivo Judicial da JFPB, ele está subordinado ao núcleo judiciário, pertencente à Secretaria Administrativa; tendo bastante autonomia mediante os demais setores da instituição, o que corrobora ainda mais sua importância dentro deste ambiente informacional.

3.1 Sobre o Arquivo Judicial

O Arquivo Judicial é considerado uma unidade informacional de médio a grande porte devido a sua grande massa documental, seu acervo é composto, exclusivamente, de processos jurídicos, importante por suas características legais e comprobatórias. O Arquivo é um setor que desempenha junto à JFPB funções relativas à guarda, organização, e identificação documental de todos os processos que foram encerrados temporariamente pela Justiça, pois a maioria dos autos ainda pode ter sua sentença recorrida na própria JFPB ou em uma instância superior. Nesse meio tempo, toda a documentação jurídica fica sob a guarda do Arquivo. Dentre as atividades específicas e diárias do setor, elencamos como as mais importantes: **o arquivamento e o desarquivamento processual, a movimentação da documentação previamente requerida para os respectivos setores da JFPB.**

Sobre o acesso e uso do conteúdo informacional na íntegra dos processos, apenas com o **desarquivamento**, tem-se acesso a todas as informações contidas em seu suporte de papel, ou seja, o acesso tangível à documentação só pode ser feito mediante a solicitação deste procedimento.

Foi através do uso dessa ferramenta de consulta que buscamos identificar os usuários do Arquivo Judicial e as necessidades informacionais por eles apresentadas, no momento da busca processual. No desarquivamento é que percebemos como os usuários do Arquivo Judicial identificam no processo as informações que mais o representa, ou seja, no momento da solicitação do desarquivamento processual os usuários do arquivo, sejam eles internos ou externos, reportam aos servidores do arquivo, quais as informações inseridas naquele documento chamaram-no maior atenção visando a recuperação documental.

4 ORGANIZAÇÃO DA INFORMAÇÃO JURÍDICA: USUÁRIOS NO CONTEXTO ARQUIVÍSTICO

Autores como Jardim e Fonseca (2004, p.3) já nos chamavam a atenção dez anos atrás sobre a importância de se explorar a temática dos estudos de usuários em arquivos já que “não se plasma uma literatura que aprofunde este tema do ponto de vista teórico e prático”. Certamente que na atual conjuntura, esse cenário vem se modificando, mas as publicações de trabalhos com a temática ainda permanecem tímidos se comparados à outras temáticas arquivísticas. Assim sendo, a ênfase da pesquisa foi nos usuários do Arquivo da JFPB e suas necessidades informacionais, além de buscar caracterizá-los por sua relação direta com a busca informacional.

Entendemos por usuários da informação todos aqueles que se deparam com algum tipo de questionamento informacional, seja nas atividades profissionais, educacionais ou simplesmente no cotidiano, e procuram sanar esses questionamentos com vistas a dar continuidade aos seus trabalhos ou simplesmente extinguir alguma dúvida. Esses usuários, em sua maioria buscam as informações em unidades informacionais, centros de documentação ou quaisquer ambientes que possuam, em sua concepção, meios para subsidiar suas indagações originadas das mais diversas áreas.

Não se pode supor que todos os arquivos possuam a mesma demanda informacional e que existam apenas usuários semelhantes. Cada arquivo possui suas especificidades e um fluxo informacional específico que deve ser identificado e trabalhado da melhor forma possível para o uso eficaz de todas as suas ferramentas, visando sempre a efetiva recuperação e uso das informações que ali estejam depositadas.

Todos os arquivos possuem suas peculiaridades informacionais, documentais e organizacionais. Portanto, enquanto unidade informacional, os arquivos precisam organizar sua massa documental de forma que a busca e a recuperação da informação sejam efetivadas com sucesso, quando e onde for solicitada pelos seus usuários. Entretanto, é preciso

considerar que “as necessidades de informação dos usuários dependem de suas atividades profissionais, área de interesse etc., enfim, variam de acordo com suas funções desempenhadas no cotidiano.” (COSTA; SILVA; RAMALHO, 2010, p.135).

Por usuário da informação, Núñez Paula (2000, p. 109, tradução nossa) assevera:

Pode se referir a uma pessoa, a um grupo ou uma entidade (usuário corporativo). Utiliza-se para designar quem utiliza a informação e seus serviços de informação. É um termo genérico e abrangente, muitas vezes o significado refere-se aos que receberam a ação dos profissionais da informação. [...]Assim, o termo usuário é relativo, e se refere a todos os que utilizam a informação, quer sejam ou não profissionais da informação.

Por usuário da informação, tem-se uma ideia clara de que toda e qualquer pessoa que numa determinada unidade informacional - com suas mais específicas características organizacionais, de gestão, suporte, descrição, acesso, uso etc – busque uma informação que em algum momento lhe falta; a busca informacional torna-se específica da unidade que o usuário acredita deter essa informação que lhe falta. Assim aparece o profissional da informação que procura assumir o papel de mediador entre o usuário e a informação que lhe falta.

4.1 Usuários do Arquivo da JFPB

Neste contexto de estudos de usuários, temos as especificidades que os caracterizam como pertencentes a uma determinada unidade informacional. O que os diferencia, muitas vezes, é justamente essa utilização de produtos e serviços que pertencem a uma unidade de forma única, como exemplo bem simplista: os livros de uma biblioteca e os documentos de um arquivo.

Entretanto, superficiais seríamos se assim definíssemos os usuários de arquivo. Estes sujeitos caracterizam-se por fazer uso do arquivo e seus documentos, mas eles fazem com que, quando considerados como premissa indispensável de funcionamento, toda a sua estrutura se modifique e se adéque de forma a atender suas demandas informacionais. O usuário é parte fundamental de uma unidade informacional e das suas diretrizes de funcionamento. Não o considerar é negligenciar a importante função daquele a quem se destina a informação que é tão trabalhosamente organizada nas unidades informacionais.

Discorrendo sobre as especificidades do usuário de arquivo, auxilia-nos o Dicionário Brasileiro de Terminologia Arquivística (2005, p.169) quando diz: “Pessoa física ou jurídica que consulta arquivos. Também chamada consulente, leitor ou pesquisador.” A existência de usuários que façam uso da informação encontrada nos arquivos, para algo mais tangível, essencial, com uso diário e não institucional ou acadêmico, por assim dizer, é pouco comum,

ou pelo menos ainda não foram relatados numa proporção tão significativa quanto na Biblioteconomia.

No Arquivo Judicial, as informações ali presentes materializadas em formato de processo contém uma diversidade informacional de extrema importância acerca da vida das pessoas. Informações estas que figuram no setor com total sigilo em respeito à instituição, mas principalmente aos usuários.

As informações processuais são organizadas de forma a representar da maneira sucinta e prática todo o teor informacional e processual. Os documentos que são formadores do acervo do Arquivo da JFPB são em demasia extensos e volumosos devido as suas características legais, assim sendo a representação de seu conteúdo, busca a utilização de descritores que sejam objetivos, claros, com teor de representatividade, de fácil acesso por parte dos usuários internos/externos e de fácil alimentação no sistema de recuperação da informação, o *Tebas*.

A etiqueta processual é o instrumento que reúne as informações mais importantes, que podem ser expostas, sobre o teor informacional e processual. Nela podemos encontrar as partes envolvidas no processo (autor e réu, seus respectivos representantes legais), o teor processual (qual o tipo de ação a ser julgada: ação ordinária, execução fiscal, inquérito policial, entre outras especificações do Direito) e o número processual que é gerado no ato de criação do processo. Apresentamos um modelo da etiqueta na figura 1:

FIGURA 1 – Modelo de etiqueta processual

PROCESSO: 0001234 – 12. 2013.4.05.8200
CLASSE: 29 AÇÃO ORDINÁRIA
AUTOR: MARIA DE FÁTIMA NUNES (PB 012345 – JOAQUIM ALMEIDA)
RÉU: JOÃO MIGUEL DE SOUZA (PB 999999 – SEM ADVOGADO)
Distribuição - Sorteio Automático em 25/04/2010 14:58
OBJETOS: 01.11.02 – Sistema Remuneratório – Servidor Público Administrativo
VARA: 3 T

Fonte: Elaborada pela autora (2014).

A etiqueta fica localizada na parte externa do processo (capa), local de fácil visualização e com a apresentação dos elementos que tem maior representatividade no teor documental: as partes envolvidas.

Os usuários do Arquivo da JFPB, internos e externos, fazem uso mesmo que de forma distinta desses elementos informacionais: o procedimento de desarquivamento (ferramenta de acesso à documentação) e a visualização da etiqueta processual (elemento com informações que representam o teor informacional do documento).

5 PROCEDIMENTOS METODOLÓGICOS: OPERACIONALIZAÇÃO DOS OBJETIVOS

Toda pesquisa possui suas características próprias, como aponta Gil (2006, p. 42), ao afirmar que a pesquisa pode ser definida “como o processo formal e sistemático de desenvolvimento do método científico”. Seu objetivo “é descobrir respostas para problemas mediante o emprego de procedimentos científicos”.

No caso específico dessa pesquisa, nossa população foi a JFPB e nossa amostra os usuários do Arquivo Judicial. Optamos por uma amostragem não probabilística chamada de amostragem por acessibilidade ou por conveniência (GIL, 2008); também encontra-se na literatura como amostragem acidental (COZBY, 2009).

Como metodologia para coleta e análise das informações, fizemos uso da Fenomenologia, objetivando “ler as entrelinhas” dos sujeitos, considerando-os em sua totalidade como ser social que possui uma vida, e esta, em sua completude, afeta diretamente todas as escolhas que este realiza.

A Fenomenologia, na perspectiva de Demo (2011, p. 250), considera que:

A subjetividade faz parte da realidade social e não pode ser acolhida metodologicamente como fator perturbante, que não deveria existir. O homem é ator, não consegue observar-se neutramente e estabelece com sua sociedade uma relação muito mais complexa que a formal-lógica da ciência clássica.

Assim, ao pensarmos numa pesquisa que tem como objetivo conhecer as necessidades de pessoas, sujeitos com grande complexidade, e todos esses elementos subjetivos que o compõem, a Fenomenologia nos apoia quanto à importância de se observar esses aspectos durante os procedimentos de coleta e análise dos dados.

Como instrumentos de coleta, optamos pela entrevista guiada e o Protocolo Verbal, sendo a entrevista utilizada com todos os sujeitos que identificamos como internos e externos ao Arquivo Judicial; e o Protocolo Verbal para registrar durante a busca informacional a interação entre o profissional do Arquivo e um usuário.

Perfilamos como sendo os sujeitos que estão “encaixados” nas competências anteriormente citadas, e que em determinado momento, fazem uso dos serviços do Arquivo Judicial, dentre estes serviços destacamos principalmente a busca informacional por processos e a solicitação de desarquivamento. Estas duas atividades exigem por parte da unidade informacional um maior e melhor desempenho na realização de suas atividades cotidianas.

Nesse sentido traçamos o perfil dos usuários do Arquivo Judicial com base nas inquietações apresentadas no momento da busca informacional; durante a solicitação de desarquivamento; e mediante as conjecturas apresentadas sobre a sua compreensão processual. Consideramos as percepções subjetivas, ou seja, a perspectiva pessoal dos sujeitos não apenas no momento da busca informacional, mas no tocante a documentação e a informação contida no documento, qual chamou mais a sua atenção enquanto representante processual.

6 DOS RESULTADOS OBTIDOS

Fazendo uso da observação participante *in loco* e também do protocolo verbal, conseguimos identificar duas grandes categorias dos usuários de arquivo na JFPB: os usuários internos à instituição e os externos a ela.

Durante a observação participante, dedicamo-nos a compreender como os usuários percebiam a informação jurídica e o porque de buscá-las na unidade informacional; quais os recursos e estratégias de busca por eles utilizados para obtenção da informação processual; o acesso e uso do SRI *Tebas*; a facilidade ou não de localização da informação tendo como parâmetro as informações disponibilizadas pela JFPB.

O protocolo verbal nos permitiu conhecer e descrever as inquietações, estratégias mentais e construção de mecanismos de busca por parte dos usuários (internos e externos), assim como seu ponto de vista sobre a organização documental e informacional, considerando claro, o acesso e uso dos instrumentos disponíveis à informação buscada (consulta processual, processos físicos, organização e representação da informação em suporte impresso e no SRI da instituição).

6.1 Usuários externos: jurisdicionados

Os usuários externos à JFPB são aquelas pessoas que em essência não conhecem bem as diretrizes de funcionamento e/ou não trabalham na Justiça; fazem parte de sua jurisdição, por terem algum vínculo processual, seja como parte (autor ou réu), como representante (advogado particular ou público), pesquisador e até mais recentemente estudante ainda em nível de graduação.

Núñez Paula (2000) corrobora nossa exposição de usuário externo. Para ele, é todo sujeito que não está vinculado à instituição, nem administrativamente e nem metodologicamente, ou seja, não existe nenhum tipo de contato com a unidade gestora que responde pela unidade informacional. O autor nos alerta ainda para a presença de usuários externos em potencial, quando os sujeitos encontram-se ligados à unidade de informação e realizam atividades como: a gestão e estratégias de organização informacional.

No ambiente jurídico é comum encontrarmos a designação de jurisdicionados para todo aquele que faz uso dos serviços da JFPB e não tem conexão trabalhista com a mesma. As NIs apresentadas pelos usuários externos estão diretamente ligadas aos processos judiciais a qual eles possuem vínculo direto, estando arquivados ou não, na JFPB. Suas necessidades partem do princípio de que **eles não detêm determinada informação sobre sua causa processual e dessa forma ficam impossibilitados de tomar algum tipo de atitude para resolver a situação estagnada** – na perspectiva pessoal de cada um – **em que se encontram**.

As motivações que os incitam a direcionar-se à JFPB - mesmo com a opção de consulta processual *online*, ou seja, com a possibilidade de busca em acesso remoto – é que: se as informações obtidas na unidade informacional forem completas o suficiente em seu ponto de vista, eles terão os meios para resolver mais facilmente a situação conflitante em que se encontram, e já estarão fisicamente na JFPB para dar mais celeridade ao que for preciso, visando assim à resolução dos seus problemas informacionais e jurídicos.

De acordo com os usuários externos, as informações processuais são de extrema importância, pois representam em suma uma “tomada de decisão”, atestam ainda a veracidade e o caráter comprobatório da documentação. Apresentaram-nos em maioria que as informações expostas na capa do processo são de fácil identificação apesar de sucintas, entretanto, os usuários externos optam por buscar a instituição porque ao consultar o SRI *Tebas*, apresenta-se apenas o resumo processual e a linguagem que nele figura não é de compreensão da maioria, mas o que podemos chamar de “juridiquês” (PASSOS; BARROS, 2009, p. 86).

6.2 Usuários internos: servidores

Os usuários internos são funcionários públicos com níveis de escolaridade médio e superior (sua grande maioria), e o ingresso à instituição foi feito mediante concurso público ou requisitados de outras instituições públicas. Os servidores das Varas Judiciais foram tipificados como sendo usuários internos dentre todos os demais servidores da JFPB, pois eles são os que mantêm uma relação constante de solicitações de serviços ao Arquivo Judicial.

Esses usuários possuem vínculo empregatício com a JFPB, mas não estão subordinados ao Arquivo Judicial.

Sobre esse tipo de usuário, especificado como interno, Núñez Paula (2000) afirma que eles devem ter uma espécie de relação administrativa ou mesmo metodológica com a unidade informacional. No caso específico dos usuários internos, estes conhecem em profundidade a documentação e todas as suas partes, entretanto, assim como os usuários externos a busca no SRI *Tebas* e seu módulo específico os deixam igualmente limitados as informações processuais. Eles conhecem as terminologias do Direito empregadas no resumo processual, não obstante, a dificuldade encontra-se em movimentar o documento e alimentar o sistema. O *Tebas* apresenta-se para eles de forma distinta, ao compararmos os módulos (Varas e Arquivo) percebemos que cada um possui características específicas relacionada ao desenvolvimento das atividades administrativas. Como não há um critério único para alimentação do sistema, ou seja, não existe uma unicidade de termos que representem a informação processual em cada setor, ocorre uma dificuldade por parte dos usuários internos em fazer uso do *Tebas* no tocante ao módulo Arquivo.

7 CONSIDERAÇÕES FINAIS

O nosso objetivo maior era o de **identificar os usuários existentes no Arquivo da JFPB e as estratégias utilizadas por eles na busca informacional**. Identificamos os usuários do Arquivo Judicial como sendo de dois tipos essencialmente por suas características de acesso e uso, tanto da unidade informacional quanto dos processos jurídicos.

Os usuários externos ou jurisdicionados são os que não possuem conhecimento prévio das funções técnico-administrativas da JFPB e tampouco o Arquivo Judicial; desconhecem em sua totalidade ou conhecem superficialmente sua rotina e funcionamento e buscam os seus serviços quando surge alguma lacuna informacional que os impossibilita à tomada de decisão e, de certa forma, os angustia.

As necessidades informacionais apresentadas pelos usuários externos mostraram-se bastante variadas, indo desde a uma simples consulta, passando por procedimentos de desarquivamento até mesmo a reabertura do processo. A linguagem utilizada para alimentar o *Tebas* e conseqüentemente aquele que aparece no momento da consulta processual, certamente é uma das maiores dificuldades descritas pelos usuários externos, para a compreensão e andamento dos seus processos, ou seja, para um efetivo uso da informação inserida no suporte impresso.

Ao fluxo informacional na perspectiva dos usuários externos apresentam-se: a elaboração de estratégias mentais; a externalização das ideias; visita à unidade informacional (momento no qual a dificuldade em entender a linguagem jurídica é descrito ao profissional de arquivo); busca processual; interação com os profissionais do Arquivo; acesso (desarquivamento) e uso do processo jurídico.

Por **usuários internos**, caracterizamos especialmente os funcionários das Varas Judiciais da JFPB que fazem uso constante dos serviços do Arquivo Judicial. Encontram-se todos os servidores, tanto das varas quanto do arquivo, subordinados à Justiça Federal, entretanto não mantêm relação de hierarquia, ou seja, cada setor responde ao seu núcleo judiciário.

Esses usuários foram tipificados como sendo internos, por estarem de alguma forma ligados ao Arquivo Judicial, seja pelas relações administrativas ou pelo desempenho de funções jurídicas, já que se encontram sob a égide da mesma instituição mantenedora; mas o contato existente como usuário interno e profissional do Arquivo ocorre efetivamente quando existe algum tipo de necessidade relacionada à ausência de informação ou a não compreensão da informação que representa o processo no *Tebas*.

As dúvidas mais frequentemente observadas ao longo da pesquisa, não apenas na aplicação dos instrumentos de coleta, eram relacionados à localização dos processos e à realização de movimentos no *Tebas*, no módulo Arquivo; para que o processo retorne não apenas fisicamente, mas também no SRI ao setor de guarda; é preciso que os servidores, mesmo sem o uso habitual do módulo Arquivo, encaminhem a documentação através dele.

Assim como fizemos com os usuários externos, incentivamos os usuários internos a exporem suas estratégias mentais (que eles elaboram mentalmente) para dirimir suas dúvidas e NIs.

No fluxo informacional, quando a localização de um processo não é frutífera, elencamos: pesquisar no *Tebas* o maior número de informações possíveis sobre o processo; consultar a localização e o resumo; quando não possuem caixa e pacote, procurar saber o ano em que foi criado e seu encerramento (o número do processo, figura nestas informações); ir até o Arquivo Judicial levando o maior número possível de informações; localizar o documento e fazer uso no desenvolvimento do seu trabalho.

Partindo da mesma ideia de expor suas estratégias mentais através do pensar alto, descrevemos como ocorre o fluxo informacional quando não se consegue utilizar o módulo arquivo: tentar movimentar através do módulo específico das Varas; selecionar quais

processos não se consegue movimentar; levar ao Arquivo; verificar algum erro na execução das tarefas; refazer com o auxílio do servidor do Arquivo Judicial.

Além dos objetivos principais da pesquisa, referindo-se especificamente aos usuários (tipos, NIs e fluxo informacional) relatamos também outro aspecto demasiado importante para a efetiva recuperação da informação, seja pelos usuários externos e/ou internos: a importância de uma representação da informação que seja uníssona em toda a JFPB, desde a descrição processual, quanto à alimentação do *Tebas*.

Ao categorizar os dois tipos de usuários da informação no Arquivo Judicial, constatamos a importância de se considerar as NIs de cada tipo, assim como a escolha terminológica que seja compreendida por ambos. Ao descrevermos os usuários e os fluxos informacionais na perspectiva de cada categoria, fez-se necessário apresentar a importância de uma representação da informação que atenda o usuário, no final do fluxo informacional (acesso e uso). A implantação do SRI *Tebas* não foi feita mediante um estudo de usuário prévio, obrigando externos e internos se adequarem ao sistema e não o oposto. Tradicionalmente chama-nos a atenção à Biblioteconomia a importância da realização desses estudos, para que quaisquer intervenções no processo de organização da informação beneficiem o usuário.

É preciso ressaltar que o uso de instrumentos de vocabulários controlados, como Tesouros, por exemplo, auxiliaria esse processo de escolha terminológica e representação da informação, tanto no suporte (processo jurídico) quanto no SRI (*Tebas*). Não podemos ignorar que a essência da instituição vem do Direito, campo do conhecimento que possui naturalmente uma linguagem singular e característica, entretanto através desse estudo percebemos que há uma diversidade de pessoas, de necessidades e de informações. Assim é preciso considerar vários aspectos ao lidar com a representação da informação em um arquivo, como por exemplo: os aspectos informacionais da instituição; respeitar os fundos arquivísticos e sua massa documental; a organização administrativa (JFPB e Arquivo); acesso e uso (com uso dos SRIs) da informação do Arquivo; e por fim não menos importante, os usuários do arquivo e suas muitas nuances e subjetividades. Seria bastante inoportuno supor que todos os usuários possuem as mesmas características, mas é assertivo prever que em algum momento há possibilidade de recuperação da informação.

REFERÊNCIAS

ARAÚJO, Carlos Alberto Ávila. Ciência da informação, biblioteconomia, arquivologia e museologia: relações institucionais e teóricas. **Enc. Bibli. R. Eletr. Bibliotecon. Ci. Inf.**, Florianópolis, . 16, n. 31, p.110-130, 2011. Disponível em:

<<https://periodicos.ufsc.br/index.php/eb/article/viewFile/1518-2924.2011v16n31p110/17765>>. Acesso em: 8 jan. 2012.

ARAÚJO, Carlos Alberto Ávila. Correntes teóricas da Arquivologia. **Enc. Bibli. R. Eletr. Bibliotecon. Ci. Inf.**, Florianópolis, v. 18, n. 37, p.61-82, mai./ago., 2013.

ARAÚJO, Carlos Alberto Ávila. Correntes teóricas da ciência da informação. **Ci. Inf.**, Brasília, DF, v. 38, n. 3, p.192-204, set./dez., 2009. Disponível em: <<http://www.scielo.br/pdf/ci/v38n3/v38n3a13.pdf>> . Acesso em: 21 jan. 2012.

ARAÚJO, Carlos Alberto Ávila. Paradigma social nos estudos de usuários da informação: abordagem interacionista. **Inf. & Soc.:Est.**, João Pessoa, v.22, n.1, p. 145-159, jan./abr. 2012. Disponível em: <<http://www.ies.ufpb.br/ojs/index.php/ies/article/view/9896/7372>>. Acesso em: 18 fev. 2013.

ARQUIVO NACIONAL. (Brasil). **Dicionário brasileiro de terminologia arquivística**. Rio de Janeiro: Arquivo Nacional, 2005. 232 p. (Publicações Técnicas, nº 51).

BORKO, H. Information science: what is it? **American Documentation**, v. 19, n.1, 1968.

BRASIL. Lei nº12.527, de 18 de novembro de 2011. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal. **Casa Civil, Subchefia para assuntos jurídicos**, Brasília, DF, 18 nov. 2011. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2011/Lei/L12527.htm>. Acesso em: 16 ago. 2013.

CAPURRO, Rafael. **Epistemologia e Ciência da Informação**. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 5, 2003. Belo Horizonte, 2003. **Anais eletrônicos...** Belo horizonte: ENANCIB, 2003. Apresentação oral. Disponível em: <http://www.capurro.de/enancib_p.htm>. Acesso em: 20 de abr. 2013.

COSTA, L. F.; SILVA, A. C. P.; RAMALHO, F. A. Para além dos estudos de uso da informação arquivística: a questão da acessibilidade. **Ci. Inf.**, Brasília, DF, v. 39 n. 2, p.129-143, maio/ago., 2010. Disponível em: <<http://dx.doi.org/10.1590/S0100-19652010000200011>>. Acesso em: 15 de jan. 2013.

COZBY, Paul. **Métodos de pesquisa em ciências do comportamento**. Tradução de Paula Inez Cunha Gomide, Emma Otta. 4. reimp. São Paulo: Atlas, 2009.

DEMO, Pedro. **Metodologia científica em ciências sociais**. 3. ed. 14 reimp. rev. e ampl. São Paulo: Atlas, 2011.

FIGUEIREDO, Nice. **Estudos de uso e usuários da informação**. Brasília: IBICT, 1994. 154p.

FREIRE, G. H.; FREIRE, I. M. **Introdução à Ciência da Informação**. João Pessoa: Editora Universitária/UFPB, 2010. 128p.

GARCIA, J. C. R. Conferências do Geórgia Institute of Technology e a Ciência da Informação: “de volta para o futuro”. **Inf. & Soc.: Estudos**, João Pessoa, v. 12, n. 1, 2002. Disponível em: <<http://www.ies.ufpb.br/ojs2/index.php/ies/article/view/153>>. Acesso em: 13 de jun. 2012.

GIL, Antônio Carlos. **Métodos e técnicas de pesquisa social**. 5. ed. São Paulo: Atlas, 2006.

GIL, Antônio Carlos. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo, Atlas, 2008.

GONZÁLEZ DE GÓMEZ, Maria Nélide. Metodologia de pesquisa no campo da Ciência da Informação. **DataGramZero - Revista de Ciência da Informação**, Rio de Janeiro, v. 1, n. 6, dez. 2000. Disponível em:

<<http://repositorio.ibict.br/bitstream/123456789/127/1/GomesDataGramZero2000.pdf>>.

Acesso em: 1 maio 2012.

JARDIM, José Maria; FONSECA, Maria Odila. Estudos de usuários em arquivos: em busca de um estado da arte. **DataGramZero – Revista de Ciência da Informação**, Rio de Janeiro, v. 5, n. 5, out. 2004. Disponível em:

<<http://www.brapci.ufpr.br/download.php?dd0=7650>>. Acesso em: 16 out. 2010.

Le COADIC, Yves-François. **A ciência da informação**. Tradução de Maria Yêda F. S. de Filgueiras Gomes. 2. ed. rev. e atual. Brasília, DF: Briquet de Lemos/Livros, 2004. 124 p.

NEVES, Dulce Amélia de Brito. **Aspectos metacognitivos na leitura do indexador**. 2004.130 f. Tese (Doutorado em ciência da informação)– Escola de ciência da informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2004.

NEVES, Dulce Amélia de Brito. Ciência da informação e cognição humana: uma abordagem do processamento da informação. **Ci. Inf.**, Brasília, v. 35, n. 1, p. 39-44, jan./abr. 2006a.

Disponível em:<<http://revista.ibict.br/ciinf/index.php/ciinf/article/viewArticle/671>>.

Acesso em: 23 mar. 2012.

NEVES, Dulce Amélia de Brito. La verbalización como registro para análisis en la investigación sobre lectura. *Anales de documentacion*, n. 9, p. 43-51, 2006b.

NEVES, Dulce Amélia de Brito; ANDRADE, Wendia O. Usuários da informação jurídica: quem são e como funciona o fluxo informacional no âmbito do Arquivo da Justiça Federal da Paraíba (JFPB). *Biblionline (João Pessoa)*, v. 8, p. 55-64, 2012.

NÚÑEZ PAULA, Isreal A. Usos y definiciones de los términos relativos a los usuarios o clientes. **Rev. Interamericana de Bibliotecología**, Medellín, Colombia, v. 23, n. 1, p.107-121, jan./ dez., 2000.

NÚÑEZ PAULA, Isreal A. Surgimento e consolidação da Documentação: subsídios para compreensão da história da Ciência da Informação no Brasil. **Perspect. ciênc. inf.**, Belo Horizonte, v.14, número especial, p. 59-79, 2009. Disponível em:

<<http://www.scielo.br/pdf/pci/v14nspe/a05v14nspe.pdf>>. Acesso em: 10 abr. 2013.

NÚÑEZ PAULA, Isreal A; LARA, M. L. G. de. A noção de documento: de Otlet aos dias de hoje. **DataGramZero - Revista de Ciência da Informação**, Rio de Janeiro, v.11 n.2 abr. 2010. Disponível em: <http://www.dgz.org.br/abr10/Art_03.htm>. Acesso em: 22. abr. 2012.

PASSOS, Edilenice; BARROS, Lucivaldo Vasconcelos. **Fontes de informação para pesquisa em direito**. Brasília, DF: Briquet de Lemos/Livros, 2009.

Paulo: Atlas, 2009.

RICHARDSON, R. J. **Pesquisa Social: métodos e técnicas**. 3. ed. rev. e ampl. São Paulo: Atlas, 2009.

IDENTIFICAÇÃO DE PRINCÍPIOS UTILITARISTAS DE JEREMY BENTHAM NA CONSTRUÇÃO DE SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO

IDENTIFICATION OF THE PRINCIPLES OF JEREMY BENTHAM'S UTILITARISM IN THE BUILDING OF KNOWLEDGE ORGANIZATION SYSTEMS

Marcos Luiz Cavalcanti de Miranda
Icléia Thiesen

Resumo: Jeremy Bentham se notabilizou como jurista por sua filosofia de cunho utilitarista desenvolvida no século das Luzes. Suas ideias influenciaram sobretudo o universo do Direito onde se enraizaram de forma inovadora, mas ganharam também outros campos que delas se apropriaram. Pouco conhecido dos profissionais da informação, o filósofo concebeu seu próprio sistema de classificação, preocupado com a definição dos conceitos jurídicos que deveriam primar pela clareza e precisão. Embora não tenham sido utilizados como foram originalmente concebidos, alguns princípios filosóficos que ele definiu nos oitocentos foram apropriados um século depois na concepção dos sistemas de organização do conhecimento construídos e ainda hoje utilizados pela Biblioteca do Congresso dos Estados Unidos. O presente trabalho tem como objetivo identificar, analisar e discutir princípios utilitaristas que influenciaram a construção de sistemas de organização do conhecimento, especificamente a *Library of Congress Classification* (LCC) e a *Library of Congress Subject Headings* (LCSH). Os procedimentos metodológicos se configuram no cotejo de aspectos da teoria utilitarista de Bentham com elementos incorporados nos referidos sistemas, permeados por informações históricas que podem explicar as influências do jurista inglês na sua concepção. Resultados preliminares indicam, como pressuposto deste trabalho, (1) aproximações entre a teoria utilitarista e os sistemas analisados, desconstruindo a ideia de que não há uma concepção teórica regendo esses sistemas, mas também (2) a necessidade de estudos mais verticais que possam evidenciar os limites e as possibilidades de aperfeiçoamento de outros sistemas de organização do conhecimento à luz da lógica benthamiana revalorizada hoje em diversos países, como a Inglaterra, sua terra natal.

Palavras-chave: Sistemas de Organização do Conhecimento. Jeremy Bentham. Classificação da Biblioteca do Congresso. Lista de Cabeçalhos de Assunto da Biblioteca do Congresso.

Abstract: Jeremy Bentham was prominent as a jurist for his philosophy of Utilitarianism developed during the Enlightenment. His ideas particularly influenced the universe of Law, where they have taken root in an innovative way, even though they were also appropriated by other fields. Little known to Information professionals, the philosopher conceived his own classification system, having in mind that the definition of juridical concepts should strive for clarity and precision. Although not used as originally designed, some philosophical principles defined by Bentham in the nineteenth century were employed a hundred years later to build the knowledge organization systems and are still used today by the Library of U.S. Congress. This study aims to identify, analyze and discuss the principles of Utilitarianism that have influenced the building of knowledge organization systems, specifically the Library of Congress Classification (LCC) and Library of Congress Subject Headings (LCSH). As a methodology, aspects of Bentham's theory and elements incorporated in the mentioned systems were collated and permeated by historical information that may explain the influences of the English jurist's conception. As a premise of this study, preliminary results indicate, (1) approaches between the theory of Utilitarianism and the analyzed systems, deconstructing the idea that no theoretical framework rules these systems, but also (2) the need of more vertical studies to show the limits and possibilities to improve other knowledge

organization systems in the light of Bentham's logic, today revalued in many countries, including England, his homeland.

Keywords: Knowledge Organization Systems. Jeremy Bentham. Library of Congress Classification. Library of Congress Subject Headings.

1 INTRODUÇÃO

A história da classificação do conhecimento acompanha par e passo a história das ciências. Ao analisarmos sua história moderna, verificamos a preeminência das ciências biológicas na definição dos caminhos seguidos desde o século XVIII - não por acaso a época das Luzes – determinando, à medida da complexificação das ciências, a necessidade crescente de uma teoria subjacente às classificações científicas.

É nessa direção que, em 1813, Augustin Pyrame de Candolle (1778-1841), botânico e médico, cunhou o termo *taxonomia*, com o sentido de teoria das classificações (BONTE; IZARD, 2007). Nomear e classificar são etapas indispensáveis sobretudo nas ciências naturais, para tornar possível um inventário das entidades em estudo a serem posteriormente catalogadas e descritas (LECOURT, 2006). Em sua obra clássica *Théorie élémentaire de la botanique*, Candolle (1813, p.23) explica a importância de uma *Teoria das classificações ou taxonomia vegetal*, título que deu à primeira parte de seu livro mais importante:

Trinta mil espécies de vegetais diferentes são conhecidas hoje na superfície do globo. Este número chegaria a mais de quarenta mil se reuníssemos tudo o que existe de inédito nas nossas coleções; e se supusermos que a Ásia, a África e a Oceania são conhecidas dos Botânicos atuais, no mesmo ponto em que a Europa o era no século dezesseis, veremos que, segundo todas as probabilidades, o globo terrestre está recoberto por mais de sessenta mil espécies de vegetais. Cada uma dessas espécies tem sua pátria, seu nome, sua forma, suas propriedades e seus usos. Todos esses conhecimentos diversos têm seus graus de utilidade (tradução e grifo nossos).

O termo grifado se repete em vários pontos do livro e não por acaso. O espírito utilitarista soprava na Europa e formava mentalidades. No plano das ideias essa filosofia teve como pai fundador Jeremy Bentham (1748-1832), jurista inglês que se debruçou sobre os estudos das instituições sociais, mas distinguindo-se especialmente no campo do Direito. Conhecedor das classificações de seu tempo criticou a forma arborescente pela qual os enciclopedistas organizaram o conhecimento, optando por seguir algumas características do sistema de organização do conhecimento do suíço C. Lineu (1707-1778), especialmente sua estrutura binária.

Botânico, zoólogo e médico, considerado o “legislador” da botânica, Lineu codificou a denominação binária de um organismo vivo, vegetal ou animal, “um nome de gênero seguido de um nome de espécie”, conforme assinalado por Hervé Le Guyader (LECOURT, 2006, p.207). Precisão e clareza constituíam o objetivo principal de sua proposta. Essa visão utilitária dos instrumentos de organização do conhecimento influenciou sobremaneira o pensador inglês. Se fizéssemos um estudo prosopográfico⁸⁰, onde diferentes personagens fossem reunidos através de suas ideias, certamente constataríamos uma circularidade cultural marcada pelo utilitarismo, sem perder de vista as nuances próprias de cada um deles. Por escapar dos objetivos do presente trabalho, adiaremos esse percurso para retomá-lo em outra oportunidade.

Jeremy Bentham se notabilizou como jurista por sua filosofia de cunho utilitarista desenvolvida no século XVIII. Suas ideias influenciaram sobretudo o universo do Direito onde se enraizaram de forma inovadora, mas ganharam também outros campos que delas se apropriaram. Tinha plena convicção de que seria possível transformar a sociedade através de leis úteis e justas que promovessem a sua segurança e, por conseguinte, a sua felicidade. Para isso seria imprescindível disciplinar o corpo social.

Pouco conhecido dos profissionais da informação, o filósofo concebeu seu próprio sistema de classificação, preocupado com a definição dos conceitos jurídicos que deveriam primar pela clareza e precisão. Embora não tenham sido utilizados como foram originalmente concebidos, alguns princípios filosóficos que ele definiu nos oitocentos foram apropriados um século depois na concepção dos sistemas de organização do conhecimento construídos e ainda hoje utilizados pela Biblioteca do Congresso dos Estados Unidos.

O presente trabalho constitui uma retomada de pesquisas anteriores⁸¹ e tem como objetivo identificar, analisar e discutir princípios utilitaristas que influenciaram a construção de sistemas de organização do conhecimento, especificamente a *Library of Congress Classification* (LCC) e a *Library of Congress Subject Headings* (LCSH). Os procedimentos metodológicos se configuram no cotejo de aspectos da teoria utilitarista de Bentham com elementos incorporados nos referidos sistemas, permeados por informações históricas e

⁸⁰ A prosopografia pode ser entendida como um método ou uma abordagem para estudos de biografias coletivas, de enfoque sociológico e histórico, buscando revelar características comuns de determinados grupos sociais. Ver Heinz (2006); Stone (1971).

⁸¹ Trata-se dos projetos de pesquisa “A informação na pré-história da Ciência da Informação: pré-conceito, natureza, episteme”, coordenado por Icléia Thiesen (2009-2012) e “Organização e representação do conhecimento em religiões de matrizes africanas”, coordenado por Marcos Miranda (2009-2012).

contextuais que podem explicar as influências do jurista inglês na sua concepção. Para fins de argumentação, outros personagens que se ligam à trajetória de Bentham são incorporados às discussões, no pressuposto de que formam elos entre si, guardadas as diferenças de abordagens das ciências e da forma de organizá-las em classificações. É o que veremos no próximo item.

2 JEREMY BENTHAM, A TEORIA UTILITARISTA E A CLASSIFICAÇÃO DO CONHECIMENTO⁸²

A compreensão do pensamento do “patriarca do utilitarismo” se vincula a diversos aspectos de sua história de vida que se refletirão nas suas propostas de reforma das instituições e nas ideias contidas em sua classificação do conhecimento, ligadas de forma indissociável.

É considerado fundador do utilitarismo clássico⁸³ que tem forte relação com o Direito. Este campo do conhecimento deve passar pelo crivo da utilidade, um de seus princípios diretores. A teoria utilitarista é uma “doutrina relativa à moral, ou seja, à questão do bem e do mal na ação humana. Foi elaborada por um conjunto de autores, principalmente britânicos, entre meados do século XVIII e do XIX” (ALLAND; RIALS, 2012, p.1498).

A utilidade é um valor e um princípio que permeia toda a sua obra. Para Bentham o homem busca em todas as coisas sua felicidade, que ele identifica com seu prazer. Mas, não se trata de uma tendência individual e egoísta, pois essa busca tem uma regra moral que o direciona ao bem estar coletivo. “O máximo de felicidade para o maior número possível de pessoas” traduz o propósito de Bentham nos planos que empreendeu para eliminar a indigência e a pobreza através da educação e do trabalho industrial nascente. (BENTHAM, 2004, p.39). Nesse aspecto, o “positivo” pode ser entendido como “útil”, por oposição à incerteza que dá lugar ao que é vago. Os utilitaristas clássicos propõem uma revisão profunda dos direitos nacionais existentes à época. Do ponto de vista formal, a ideia central desta corrente é a codificação que ocupa uma parte determinante na reforma de Bentham, mas

⁸² Não nos propusemos a descrever detalhadamente as ideias de Bentham, mas a evidenciar os princípios de sua filosofia que estão refletidos nos seus projetos e, por conseguinte, na forma como concebe o conhecimento. Para um estudo da sua obra ver, entre outros, Brunon-Ernst (2007); Cléro & Laval (2002); Cléro (2006); Centre Bentham (<http://www.centrebentham.fr>); *Bentham Project* (<http://www.ucl.ac.uk/bentham-project/>). Os pesquisadores do *Centre Bentham* dispõem de um sítio na Internet, promovem colóquios internacionais sobre o tema, publicam online a *Révue d'études benthamiennes*, além de empreenderem a tradução da obra de Bentham para a língua francesa. A atualidade de Bentham é incontestável.

⁸³ Para muitos autores esse ramo da filosofia constitui uma migração do positivismo francês de Comte para a Inglaterra.

aprovada também por John Austin, James Mill, seu filho John Stuart Mill (ALLAND; RIALS, 2012, p.1501).

O personagem aqui referido recebeu uma educação austera que marcará toda a sua vida. Apesar de pertencer a uma família bem aquinhoadada em termos de recursos financeiros, é educado em colégio interno, dispondo de pouco dinheiro para se manter no dia a dia, o que o colocava em situação desigual à dos colegas. Michelle Perrot explica que ele “tem relações tensas com seus companheiros” (PERROT, 2000, p.113). Essa experiência em uma instituição fechada, com forte característica disciplinar terá reflexos, ainda, no modelo de escola chrestomáthica⁸⁴ que o jurista idealizará já na idade madura. Todo seu projeto reflete as bases da filosofia utilitarista, doutrina cujo princípio “aprova ou desaprova uma ação, qualquer que seja, segundo sua tendência de aumentar ou de diminuir a felicidade da parte cujo interesse está em questão” (CLÉRO; LAVAL, 2002, p.54). Pode-se resumi-la na concepção do panoptismo⁸⁵ que se faz refletir na escola idealizada, na deontologia, no código constitucional, no sistema de penas e recompensas, etc. Este último servirá de base para a mensuração do que útil e positivo para a sociedade. Seu objetivo era tornar o conhecimento mais adequado às necessidades sociais através do ensino da população inglesa face às mudanças advindas da industrialização nascente e do pensamento iluminista. Nesse aspecto, o trabalho constitui um valor fundamental. Bentham anteviu a proximidade e expansão das novas formas de trabalho industrial que pressupunham mudanças qualitativas na formação da mão de obra dos novos trabalhadores.

Aos doze anos de idade torna-se estudante de Direito em Oxford, conforme a vontade de seu pai, ocasião em que se decepçiona com o ensino ali ministrado, ressaltando a

⁸⁴ *Chrestomathia* é o título da obra de Bentham. O termo de sua lavra significa “o que conduz à aprendizagem útil”, obra que contém seu amplo esquema de pensamento e que engloba seu esquema de classificação do conhecimento. Os originais datam de 1816-1817 e foi editada em 1843, em Edimburgo, por John Bowring (1792-1872), negociante e advogado de Londres que se tornou seu amigo e, “após a sua morte, seu executor testamentário, encarregado da publicação das obras do mestre” (BRUNON-ERNST, 2007, p. 233). A tradução francesa realizada por Jean-Pierre Cléro, publicada em 2004 e analisada neste trabalho, considerou a 1ª tradução inglesa, mas baseou-se principalmente na edição inglesa de 1983, cf. explicações de seu tradutor (CLÉRO, 2004, p.LXXI. In: BENTHAM, 2004). O mesmo autor explica que a tradução do Apêndice IV dessa obra foi realizada por George Bentham, sobrinho do jurista e filho de Samuel Bentham.

⁸⁵ Termo dos mais conhecidos do vocabulário de Bentham constitui ao mesmo tempo um projeto amplo para a construção de prisões, casas de indústria e uma escola, pois o princípio da utilidade requer um projeto de educação. Mas, segundo Cléro & Laval (2002, p.19), constitui também um princípio de inspeção, sem o qual não seria possível compreender o pensamento utilitarista de Bentham. Devemos a Michel Foucault a popularização do termo e do projeto panóptico, especialmente em *Vigiar e punir: o nascimento das prisões* (1993) e em *A verdade e as formas jurídicas* (1996).

depravação moral que reinava naquele lugar destinado a formar o clero anglicano, o que poderia de certa forma explicar seu ateísmo declarado na maturidade (BRUNON-ERNST, 2007, p.30; AUTORES, 2009, p.224). Aos quinze anos volta a Londres e ingressa nos estudos jurídicos ministrados no Lincoln's Inn, "uma espécie de faculdade de estudos jurídicos muito fechada" (PERROT, 2000, p.113), repetindo a mesma decepção com o sistema vigente, o que o levará mais tarde a pensar na reforma das instituições⁸⁶ e a jamais exercer a advocacia. É nesse momento de sua vida, ainda muito jovem, que tomará conhecimento da obra de Lineu, considerado um dos mais expressivos taxonomistas de seu tempo (BURKE, 2003, p.79). Lineu utiliza a bifurcação e a bipartição para classificar o conhecimento da botânica. Bentham fará o mesmo procedimento binário ao refazer o sistema classificatório arborescente da arte e da ciência feitos por Diderot e d'Alembert na Enciclopédia (BENTHAM, 2004). O jurista se refere inúmeras vezes a essa "obra célebre", em sua *Chrestomathia*. Em «La carte encyclopédique de D'Alembert ou l'esquisse en forme de table: ses imperfections», por exemplo, Bentham relaciona as falhas ali encontradas:

1º O tema mesmo da obra, designado inadequadamente. 2º A *fonte primeira* das *divisões* escolhida de forma inoportuna. 3º O *esquema de divisão*, solto e irregular. 4º As *chamadas*, em numerosos casos [são] inapropriadas. 5º As *distinções infundadas*, em inúmeros casos de princípios, inapropriadas. 6º *Repetições frequentes*: o mesmo objeto repetido inúmeras vezes sob diversos nomes. 7º A *textura incompleta* do discurso: sem *verbos* e, conseqüentemente, sem *proposições*; nada além de *substantivos*, com isso e aquilo, um *artigo* ou um *adjetivo* (BENTHAM, 2004, p.179-180).

A utilização da estrutura por bifurcação exaustiva é assim explicada e justificada por Bentham:

a qualidade exaustiva será sempre útil na divisão de todo conjunto lógico; e, para que a divisão forneça a prova dessa qualidade, é preciso que seja sempre redigida a partir dos princípios que acabo de estabelecer. Em qualquer que seja o ramo, quanto mais continuar essa divisão, mais se obterá vantagens pela precisão que ela dará às ideias que se concebe do assunto (BENTHAM, 2004, p. 206).

As concepções de educação do jurista sofreram diversas influências, como a de John Locke e Hobbes (BRUNON-ERNST, 2007, p.215). Essas ideias aparecem no seu sistema de penas e recompensas, tendo no horizonte a felicidade esperada e conquistada por e para todos. Outro ponto nevrálgico de seu sistema de pensamento é a importância dada à linguagem,

⁸⁶ Bentham projetou "uma prisão nos anos 1787-1791, segundo o modelo panóptico, uma Casa de Indústria no período de 1796-1798 e um externato nos anos 1816-1817" (BRUNON-ERNST, 2007, p.17). A concepção de tais projetos, todos direcionados à criação de instituições, obedeceu ao processo de maturação de ideias utilitaristas. (THIESEN; MIRANDA, 2009)

preocupação também demonstrada por outros filósofos seus contemporâneos, como Hume, Locke e Berkeley. Nesse sentido, Cléro & Laval afirmam que “o utilitarismo é fundamentalmente uma teoria da linguagem” (2002, p.7). Para Bentham “as leis só podem ser feitas com palavras. A vida, a liberdade, a propriedade, a honra, tudo o que nos é caro depende da escolha das palavras” (CLÉRO; LAVAL, 2002, p.8).

Veremos a seguir de que maneira as ideias do jurista se refletiram (ou não) na classificação do conhecimento proposta em *Chrestomathia*⁸⁷. Nessa obra Jeremy Bentham apresenta o conteúdo da instrução chrestomáthica (tabelas I e II), reproduz (tabela III) e critica a estruturação do conhecimento de Diderot e d’Alembert segundo suas convicções, assim como faz emendas à Árvore de Porfírio (Tabela IV).

Por fim, na Tabela V estrutura sua classificação do conhecimento segundo o modo de bifurcação, tal como pensou as Artes e as Ciências (THIESEN; MIRANDA, 2009, p.222). No mundo do Direito a imprecisão e a ambiguidade presentes no vocabulário jurídico podem acarretar danos e prejuízos à sociedade. É nesse sentido que o jurista apresenta a linguagem como representação e comunicação do conhecimento e da informação, conforme a mencionada Tabela V:

⁸⁷ Não é possível apresentar a totalidade da Tabela V sem perda de informação que decorreria de sua inserção nos limites deste trabalho, razão pela qual apresentamos um resumo esquemático para fins de argumentação.

ESQUEMA 1: Table Encyclopédique ou Table de l'Art et de la Science.

Eudémonique
Coenoscopique
Idioscopique
Ontologie Idioscopique
Somatoscopique
Somatique
Pososcopique
Mathématique
Alegomorphique
Mathématique Alemorphique
Morphoscopique
Poïoscopique
Somatique Poïoscopique
Physiurgoscopique
Histoire Naturelle
Uranoscopique
Épigéoscopique
Physiurgie Épigéoscopique
Abioscopique
Embioscopique
Physiurgie Embioscopique
Azooscopique
Zooscopique
Anthropurgoscopique
Philosophie Naturelle
Coenoscopique
Mècanique
Idioscopique
Chimie
Mècanique et la Chimie
Anapirique
Catastaticochrestique
Pneumatoscopique
Pneumatologie
Nooscopique [<i>Connaissances</i>]
Pneumatologie Nooscopique
Plasioscopique
Coenonesioscopique [<i>Communication</i>]
Nooscopie Cenonesioscopique
Didactique [<i>Information</i>]
Pathocinetique
Pathoscopique
Pneumatologie Pathologique
Ergastique
Anergastique

Fonte: BENTHAM, J. (2004, [1843]); THIESEN; MIRANDA (2009)

Em *Chrestomathia* Bentham discute a melhor forma de classificar o conhecimento com vistas a evitar ambiguidades, justificando a opção pela bifurcação e a exaustividade.

Deve-se minuciosamente na definição de cada termo inserido no esquema proposto. O jurista inventou diversos vocábulos de origem grega, muitos dos quais constituíam neologismos. Alguns foram posteriormente dicionarizados e outros permaneceram obscuros. Ao cunhar uma terminologia própria ao pensamento filosófico utilitarista, Bentham se diferencia de seus contemporâneos pela univocidade do sentido que os termos carregam – intenção coerente com seus propósitos, mas simultaneamente pouco difundidos fora do universo jurídico. Percorreu, portanto, o caminho inverso das práticas classificatórias se considerarmos que o conhecimento representado nos sistemas de organização é aquele aceito e consolidado pela área à qual se destina.

3 SISTEMAS DE ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO DA LIBRARY OF CONGRESS

O conceito de “Sistemas de Organização do Conhecimento” (SOC) compreende todos os tipos de instrumentos utilizados para organizar a informação e promover o gerenciamento do conhecimento, incluindo os esquemas de classificação – que organizam materiais em nível geral (como livros em uma estante); cabeçalhos de assunto – que proveem o acesso mais detalhado; catálogos de autoridade – que controlam versões variantes de informação fundamental (como nomes geográficos e nomes pessoais) e outros instrumentos menos tradicionais - como redes semânticas e ontologias. Os sistemas de organização de conhecimento são mecanismos para organizar a informação e constituem o coração de todo Sistema de Recuperação de Informação, como biblioteca, museu e arquivo (HODGE, 2000).

A decisão de usar um sistema de organização do conhecimento é central ao desenvolvimento de qualquer SRI, sendo a classificação a principal atividade neste processo. Como sabemos, o processo de classificar é inerente aos seres humanos desde cedo, em contato com o mundo através da comparação entre objetos novos ou da experiência com aqueles com que estamos familiarizados, identificando padrões e categorizando o que é novo em nosso quadro de referência já existente. A ênfase no desenvolvimento de SOC pode ser encontrada nos escritos de nossos mais antigos filósofos, muitos dos quais continuam influenciando nossa visão do mundo. Por exemplo, o esforço de Aristóteles para categorizar o conhecimento em grupos (como física, política ou metafísica) está refletido em nosso idioma, nossa educação e nossa ciência (MIRANDA et. al., 2012)

O esquema de classificação original da Biblioteca do Congresso, usado entre 1800 e 1814, estava baseado nos trabalhos filosóficos de Francis Bacon que herdou da tradição inglesa, que por sua vez originou a CDD e esta a CDU. Começando em 1814, a influência de

Thomas Jefferson pode ser vista na coleção da Biblioteca do Congresso. Foi Jefferson quem reclassificou a biblioteca, à luz de uma filosofia mais humanista (LESK, 1997).

Lesk afirma que não existe um único SOC com o qual todos concordem, mas especula que um único seria vantajoso, porém seria improvável que tal sistema fosse desenvolvido. A questão cultural pode limitar um sistema de organização do conhecimento de forma que o que é significativo para uma cultura pode não ser necessariamente para outra (LESK, 1997). Então, habitamos um mundo de múltiplas visões, com várias maneiras para organizar o conhecimento. Até porque toda classificação, toda organização, pressupõe uma escolha, um corte epistemológico frente a seu objetivo específico. Apesar dessa diversidade, os SOC possuem as seguintes características comuns que são críticas em relação ao seu uso para organização do conhecimento em qualquer ambiente de informação: impõem uma visão particular de mundo sobre uma coleção e sobre os itens que a compõem; a mesma entidade pode ser caracterizada de modos diferentes, dependendo do SOC que é utilizado.

Beghtol (2002) introduz o princípio da garantia cultural que reside na ideia de um sistema de organização do conhecimento ser o mais apropriado possível e útil para os indivíduos de uma cultura apenas se ele for baseado nas suposições, valores e predisposições dessa mesma cultura. Inversamente, se um sistema não é baseado nessas suposições, ele será apropriado e útil a uma menor extensão de indivíduos na cultura. E apresenta como solução o uso deste princípio de maneira que as notações - utilizadas nos esquemas de classificação para representar e organizar o conhecimento - tenham a habilidade de “admitir novos conceitos apropriadamente e acomodá-los nas relações corretas com outros conceitos” (BEGHTOL, 2002, p.35). O princípio da hospitalidade cultural, portanto, complementa o princípio da garantia cultural.

3.1 A Library of Congress Classification

A *Library of Congress* (LC) foi criada por um ato do Congresso em 1800 quando o então Presidente John Adams assinou termo para a transferência da sede do governo da Filadélfia para a nova capital na cidade de Washington e foi inaugurada com uma coleção de 3.000 volumes. Os livros, que antes eram ordenados por tamanho, em 1892 já estavam divididos em 18 classes, baseadas nas classificações de Francis Bacon, com adaptação de Diderot e d’Alembert. Em 1815 fora adquirida a coleção de Thomas Jefferson, constituindo assim a nova biblioteca (PIEDADE, 1983).

Após a mudança de prédio, em 1897, os bibliotecários sentiram a necessidade de criar um novo sistema de classificação que comportasse o crescente acervo. Designados por John

Russel Young, então diretor da entidade, James Hanson e Charles Martel tomaram por guia a Classificação Expansiva de Cutter, “introduzindo grandes modificações, especialmente quanto à notação” (PIEDADE, 1983, p.153). A partir deste planejamento em linhas gerais, cada classe foi entregue a diversos especialistas, derivando daí as pequenas diferenças que ocorrem de uma classe para outra.

A Classificação da *Library of Congress* baseou-se em 21 classes principais representadas de A-Z, exceto pelas letras I, O, N, X e Y, deixadas para futuras expansões, sendo igualmente adotada por diversas bibliotecas dos EUA e no mundo. O critério da previsibilidade científica está presente também no pensamento utilitarista e na classificação de Jeremy Bentham.

ESQUEMA 2: As 21 classes da *Library of Congress Classification*.

Classe	Assuntos cobertos
A	<i>General Works</i>
B	<i>Philosophy, Psychology, and Religion</i>
C	<i>Auxiliary Sciences of History</i>
D	<i>General and Old World History</i>
E	<i>History of America</i>
F	<i>History of the United States and British, Dutch, French, and Latin America</i>
G	<i>Geography, Anthropology, and Recreation</i>
H	<i>Social Sciences</i>
J	<i>Political Science</i>
K	<i>Law</i>
L	<i>Education</i>
M	<i>Music</i>
N	<i>Fine Arts</i>
P	<i>Language and Literature</i>
Q	<i>Science</i>
R	<i>Medicine</i>
S	<i>Agriculture</i>
T	<i>Technology</i>
U	<i>Military Science</i>
V	<i>Naval Science</i>
Z	<i>Bibliography, Library Science, and General Information Resources</i>

As classes são publicadas independentemente umas das outras e cada uma tem seu próprio índice, sofrendo revisões e acréscimos, conforme a expansão do acervo, publicadas quadrimestralmente no *L.C. Classification: Addition and changes*.

A extensão de cada classe, sua maior ou menor especificidade depende em grande parte do desenvolvimento da coleção, portanto o desenvolvimento do sistema baseia-se na garantia literária.

Sayers (1962, p.151) considera a LCC um sistema utilitário, sem base filosófica, mas, sendo originário do conhecimento escrito e sendo este uma expressão da ordem natural, procurada pelas classificações filosóficas. Para Foskett (1985, p. 112) "esta distinção é superficial, desde que só o conhecimento expresso pode ser classificado e quanto mais próximo está na ordem da natureza maior é sua verdade objetiva e seu sucesso em aumentar nosso poder sobre a natureza".

A LCC é o mais importante sistema utilitário, sem apoio em bases científicas de que se tem conhecimento. Embora a arrumação dos assuntos dentro das classes seja inteiramente arbitrária, dependendo das necessidades da própria Biblioteca, ela tem sido, desde sua criação, em fins do século XIX, até hoje um sistema atuante, em constante expansão (BARBOSA, 1969, p.53).

O sistema da Biblioteca do Congresso é muito detalhado, podendo ser considerado como uma série de classificações especializadas, bastante enumerativo, mas que recorre à síntese, quando aplica suas inúmeras tabelas auxiliares.

Na avaliação de Langridge (1977, p. 87-88) a LCC é o menos sistemático de todos os esquemas de classificação, com possibilidade frequente de classificação-cruzada. Produto de uma escola prática obstinada que parece pensar que teoria não tem nenhuma relação com prática. Como o esquema de Dewey, é frequentemente defendida pelo comentário "funciona". Desde que não se trata de fato de um esquema funcionar ou não funcionar, mas sim de funcionar com grande grau de eficiência, o comentário não tem muita significação.

O sistema da Biblioteca do Congresso tem a flexibilidade para classificar qualquer tipo de material, é muito detalhado, enumerativo, porém recorrente à síntese, quando aplicadas suas inúmeras tabelas auxiliares. É um esquema prático para aqueles que acreditam em soluções simples e, atualmente, apresenta como índice geral a *Library of Congress Subject Headings*, como veremos a seguir.

3.2 A Library of Congress Subject Headings

A *Library of Congress Subject Headings* surgiu por obra de seu ilustre bibliotecário Charles Ammi Cutter, no final do século XIX, quando a biblioteca foi transferida para sua nova sede.

Para a Biblioteca do Congresso Cutter preparou um *catálogo-dicionário* no qual eram listados, juntos, em ordem alfabética, autores, títulos e assuntos dos livros (GNOLI; VITTORIO; ROSATI, 2006, p. 36). Este tipo de catálogo ainda é utilizado atualmente, sobretudo em pequenas bibliotecas públicas. Com o tempo, os termos do catálogo foram sendo refinados e formalizados, enriquecendo-se de *remissivas* para as formas preferidas de um termo ou para suas variantes. Deste modo, quando os usuários procuram um termo não utilizado pelo catálogo são então reenviados ao termo corrente. Na verdade, enquanto a identificação de autores e de títulos na maior parte dos casos é bastante imediata, para os assuntos é necessário seguir algumas regras convencionais para se assegurar que:

- um conceito seja representado sempre com a mesma forma (uniformidade);
- a mesma forma represente sempre o mesmo conceito (univocidade);
- o tema de cada documento seja descrito de modo completo (exaustividade);
- um conceito seja representado exatamente na linguagem de indexação (especificidade);
- o catálogo responda a todas as expectativas dos usuários (previsibilidade).

Em 1876 Cutter elaborou as *Rules for a dictionary catalog*, onde estabeleceu um conjunto genérico de regras para a construção e o arranjo de cabeçalhos de assunto, resumidas em três princípios básicos : 1º - Princípio específico; 2º - Princípio do uso e 3º - Princípio sindético. No primeiro os assuntos devem dar entrada pelo termo mais específico e não pela classe que está subordinada; o segundo preconiza que deve haver compatibilidade entre a terminologia e a forma que os usuários utilizam a informação; o terceiro pressupõe a ligação de assuntos correlacionados através de uma rede de referências cruzadas.

The List of Subject Headings for the use in Dictionary Catalogs preparada por um comitê da *American Library Association* (ALA) e publicado em 1895, foi usado como base para a elaboração de cabeçalhos de assunto da Biblioteca do Congresso.

Deste modo, progressivamente se consolidaram os *Library of Congress Subject Headings* (LCSH), tornando-se logo os *cabeçalhos* mais importantes do mundo, atualmente em sua 36ª edição. Do mesmo modo que seu esquema de classificação, a LCSH se difundiu graças à distribuição das fichas da Biblioteca do Congresso para outras bibliotecas, depois em forma de fitas magnéticas e atualmente se encontram em formato digital.

No que se refere à ordem sistemática, a alfabética apresenta pelo menos duas vantagens evidentes: é de compreensão imediata dos usuários e possibilita a inserção a qualquer momento (e lugar) de novos cabeçalhos, constituindo a chamada *hospitalidade*. Por outro lado, temas afins descritos por diversos termos, como antílopes e zebras, encontram-se distantes sem qualquer regra sistemática (GNOLI; VITTORIO; ROSATI, 2006, p. 42). Numa ordem sistemática esses assuntos estariam agrupados na classe de mamíferos, numa ordenação alfabética estariam dispersos.

Lista de cabeçalho de assunto é um tipo de sistema de organização do conhecimento, assim como os esquemas de classificação, os tesauros, as taxonomias e as ontologias e tem como finalidade a representação dos assuntos dos documentos por cabeçalhos de assunto que arrola, exercendo certo controle sobre os termos de indexação a serem adotados pelo bibliotecário numa determinada unidade de informação.

Uma lista de cabeçalho de assunto é composta pelas instruções de uso, pelos cabeçalhos de assunto arrolados, pelas remissivas “ver” e pela estrutura sindética (rede de remissivas “ver também”). Constitui parte essencial de um sistema de recuperação da informação, pois reflete diretamente na recuperação dos assuntos ordenados no catálogo alfabético de assunto ou nas bases de dados.

A *Library of Congress Subject Headings* (LCSH) encontra-se na sua 36ª edição (2014, em 6 volumes), continuando um trabalho cumulativo de cabeçalhos de assunto estabelecidos pela Biblioteca do Congresso dos Estados Unidos da América (LC) desde 1898. A primeira edição da LCSH, denominada *Subject Headings used in the Dictionary Catalogues of the Library of Congress*, foi impressa em partes entre 1909 e 1914. Requisitou-se a edição de listas suplementares, seguida de uma 2ª edição em 1919. A 4ª edição (1943) da LCSH introduziu o conceito de *pattern-headings*, significando que determinados cabeçalhos são usados como modelos para subdivisões que podem ser usadas com outros cabeçalhos na mesma categoria de assunto. Edições posteriores foram publicadas em intervalos irregulares.

Até 1974, catalogadores de assunto da LC regularmente estabeleciam combinações de subcabeçalhos, submetendo-os para revisão, aprovação e impressão na LCSH. Em 1974, também ficou decidido que a maioria das combinações dos subcabeçalhos ficariam, no futuro, elaborados em conformidade com regras, ao invés de autorizações individuais. O título mudou para *Library of Congress Subject Headings* (LCSH) quando foi publicada a 8ª edição em 1975. Sua longa introdução incluía uma lista das subdivisões mais usadas (*free-floating subdivisions*) e notas de escopo orientando seu uso.

O fato de a lista ter se expandido através dos tempos reflete a variedade de princípios de centenas de catalogadores que têm contribuído na elaboração, inclusão, alteração e/ou exclusão dos cabeçalhos de assunto.

A LCSH também teve sua estrutura modificada a partir da 11ª edição em 1987. Até a 10ª edição utiliza-se para apresentação dos cabeçalhos de assunto os seguintes sinais: sa - (*see also*) - indicava uma referência para um tópico relacionado ou subordinado; x - (*see from*) - indicava uma referência de uma expressão não adotada como subcabeçalho e xx - (*see also from*) - indicava um cabeçalho relacionado ou genérico a qual uma referência ver também é feita. A partir da 11ª edição, depois de várias reavaliações feitas pelo comitê da Biblioteca do Congresso, decidiu-se pela estrutura semelhante a do tesouro:

UF - *used for*-----x

BT - *broader topic*-----xx

RT - *related topic*-----xx/sa

NT - *narrower topic*-----sa

A 36ª edição da Lista de Cabeçalho de Assunto da Biblioteca do Congresso (LCSH 36) contém cabeçalhos estabelecidos pela Biblioteca até junho de 2014. Os cabeçalhos de assunto incluídos nesta lista foram obtidos através da criação de um banco de dados constituído por todos os registros de cabeçalhos de assunto e de subdivisões pesquisados no catálogo de autoridade de assunto, na Biblioteca do Congresso. Cerca de 336.090 registros de autoridade constavam então no catálogo.

Na LCSH36 o conhecimento está organizado em cabeçalhos e subcabeçalhos de assunto ordenados alfabeticamente, constituindo um total de aproximadamente 336.090 cabeçalhos de assunto, ou seja, 9.000 a mais que a 35ª edição.

O banco de dados de autoridade de assunto no qual os cabeçalhos de assunto nesta edição foram estabelecidos indica que foram arrolados cerca de 23.844 cabeçalhos de nomes pessoais, dos quais 22.686 representam cabeçalhos de nomes de família, 9.327 cabeçalhos corporativos, 9 cabeçalhos de reuniões ou conferências, 487 títulos uniformes, 238.859 cabeçalhos de assuntos tópicos e 59.956 cabeçalhos de assuntos geográficos. Há 769 referências USE, 4.352 referências VER TAMBÉM, 293.125 referências de uma rubrica para outra, utilizável e 352.269 referências de cabeçalhos de assunto não permitidos para cabeçalhos de assunto permitidos.

A criação e revisão de cabeçalhos de assunto é um processo contínuo. Aproximadamente 5.000 novos cabeçalhos, incluindo aqueles com subdivisões, são adicionados a LCSH a cada ano. As propostas de novos cabeçalhos e revisões existentes são

submetidas pelos catalogadores na Biblioteca do Congresso e aos participantes do *Subject Authority Cooperative Program* (SACO). As propostas aprovadas passam a integrar o catálogo de autoridade de cabeçalhos de assunto em linha, na Biblioteca do Congresso, a partir do qual várias publicações são criadas.

3.3 Ressonâncias utilitaristas nos Sistemas de Organização do Conhecimento da *Library of Congress*

Alguns aspectos da teoria utilitarista de Bentham podem ser cotejados com elementos incorporados nos sistemas de organização do conhecimento da *Library of Congress*. A LCC é considerada como o maior sistema de organização do conhecimento utilitário em uso. Foi elaborada, é desenvolvida e atualizada de acordo com as necessidades da própria Biblioteca do Congresso dos EUA. Para Mann (1950, p.100) a LCC "é essencialmente uma classificação de livros porque os livros em cada classe já tinham sido arrumados e estudados antes do esquema ser feito. Provavelmente nunca haverá melhor oportunidade para se criar uma classificação baseada na própria utilização dos livros."

A flexibilidade e a expansibilidade praticamente ilimitada da LCC, a variedade de artifícios de notação para a subdivisão de assuntos por forma, por local, dentre outros permitem não só a adição e incorporação de novos assuntos, mas tornam possível com um símbolo de três algarismos subdividir gradualmente classe por classe.

Como no sistema contido na obra *Chrestomathia* de Jeremy Bentham, a LCC apresenta a estruturação do conhecimento com base na Classificação Expansiva de Cutter, que por sua vez se baseou em Diderot e d'Alembert organizando o conhecimento segundo o modo de bifurcação.

Da mesma maneira, encontramos correspondência entre a teoria geral da classificação e a teoria utilitarista, como por exemplo, para a sistematização de cabeçalhos de assunto com base nos princípios da uniformidade, da univocidade, da exaustividade, da especificidade, da previsibilidade. Assim como para a teoria de Cutter para a construção e o arranjo de cabeçalhos de assunto: princípio específico; princípio do uso e princípio sindético.

Aprofundando mais nossos estudos encontramos também algumas semelhanças na teoria da classificação de Bliss (1939) e na teoria da classificação de Ranganathan (1967), como o princípio da sequência útil para a construção de renques⁸⁸.

4 CONSIDERAÇÕES FINAIS

⁸⁸ Renques são classes de termos formadas a partir de uma única característica de divisão.

Os resultados preliminares obtidos até aqui indicam, como ponto de partida deste trabalho, diversas aproximações entre a teoria utilitarista de Jeremy Bentham e os sistemas de organização do conhecimento da Library of Congress Classification e Library of Congress Subject Headings aqui analisados.

Neste sentido, fornecemos alguns subsídios que podem contribuir para a desconstrução da ideia de que não há uma concepção teórica regendo esses sistemas. Além disso, permanece a necessidade de estudos mais verticais que possam evidenciar os limites e as possibilidades de aperfeiçoamento de outros sistemas de organização do conhecimento à luz da lógica benthamiana revalorizada hoje em diversos países, como a Inglaterra, sua terra natal. O utilitarismo clássico, que tem em Jeremy Bentham seu maior representante, como vimos, é hoje um campo de estudos fértil sobre o qual se debruçam pesquisadores de diversas áreas do conhecimento da França e da Inglaterra.

O que resta hoje de seu sistema de pensamento? Entre as diversas contribuições que se acham registradas na história das classificações – um capítulo que merece ainda muitas pesquisas – suas ideias são exploradas e analisadas nos campos da filosofia, da economia, da história, do direito, da sociologia, entre outros. A existência de uma revista publicada em inglês e francês, desde 2006 – *Révue d'études benthamiennes*, internacional e pluridisciplinar - atesta sua importância nesses dias em que a informação e o conhecimento se expandem de modo expressivo, num mundo cada vez mais voltado para a busca da comunicação em tempo real, o que pressupõe novas formas de organizar o conhecimento. Além desses aspectos, estudos comparativos entre as diversas classificações podem ser direcionados à compreensão de seus formuladores, com o objetivo de identificar comunidades de sentido que expliquem a coincidência de características que migram de um sistema a outro, guardando, entretanto, suas próprias identidades.

REFERÊNCIAS

ALLAND, D.; RIALS, S. (Dir.). **Dictionnaire de la culture juridique**. Paris: Quadrige / LAMY-PUF, 2012.

BARBOSA, A.P. **Teoria e prática dos sistemas de classificação bibliográfica**. Rio de Janeiro: Instituto Brasileiro de Bibliografia e Documentação, 1969.

BEGHTOL, C. Universal concepts, cultural warrant, and cultural hospitality. In: LOPEZ-HUERTAS, M. J. (Ed.). **Challenges in knowledge representation and organization for the 21st century**: integration of knowledge across boundaries. Würzburg: Ergon-Verlag, 2002. p. 45-49.

BENTHAM Project. University College London (UCL). Disponível em: <<http://www.ucl.ac.uk/>>.

BENTHAM, J. *Chrestomathia* [1816-1817]. In: SMITH, M. J.; BURSTON, W. H. (Ed). **The collected works of Jeremy Bentham**, Oxford, Clarendon Press, 1983.

BENTHAM, J. **Chrestomathia**: recueil de feuillets qui expliquent le projet d'une institution destinée à être mise en œuvre sous le nom de Externat Chrestomathique ou École Chrestomathique pour l'extension du nouveau système d'instruction aux branches supérieures de l'enseignement, à l'usage des rangs moyens et supérieurs de la vie. Traduction, notes et introduction de Jean-Pierre Cléro. Paris: Cahiers de l'Unebévue, 2004.

BLISS, H.E. **The organization of knowledge in libraries**. New York : H. W. Wilson, 1939.

BONTE, P.; IZARD, M. (Dir.). **Dictionnaire de l'ethnologie et de l'anthropologie**. Paris: Quadriga / PUF, 2007. p.153

BRUNON-ERNST, A. **Le panoptique des pauvres**: Bentham et la réforme de l'assistance en Angleterre. Paris: Presses Sorbonne Nouvelle, 2007.

BURKE, P. **Uma história social do conhecimento**: de Gutenberg a Diderot. Rio de Janeiro: Jorge Zahar Editor, 2003.

CENTRE BENTHAM: traduire, commenter, diffuser l'œuvre de Jeremy Bentham. Disponível em: <<http://www.centrebentham.fr>>.

CLÉRO, J-P. **Bentham philosophe de l'utilité**. Paris: Ellipses, 2006.

CLÉRO, J-P.; LAVAL, C. **Le vocabulaire de Bentham**. Paris: Ellipses, 2002.

CONDOLLE, A.P. de. **Théorie élémentaire de la botanique** ou Exposition des principes de la classification naturelle et de l'art de décrire et d'étudier les végétaux. Paris: chez Déterville, Libraire, rue des Hautefeuilles, n.8, 1813. Disponible en : http://books.google.com.br/books?id=ThMAAAAQAAJ&printsec=frontcover&hl=pt-PT&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false Accès en: 21 de julho de 2014.

FOSKETT, D. J. **Subject and information analysis**. New York: Marcel Dekker, 1985.

FOUCAULT, M. **A verdade e as formas jurídicas**. Rio de Janeiro: NAU, 1996.

FOUCAULT, M. **Vigiar e punir: o nascimento das prisões**. Trad. de Lígia M. Ponde Vassallo. 10. ed. Petrópolis (RJ): Vozes, 1993.

GNOLI, C.; VITTORIO, M.; ROSATI, L. **Organizzare la conoscenza**. Dalle biblioteche all'architettura dell'informazione. Roma: Tecniche nuove, 2006.

HEINZ, F.M. (Org.). **Por outra história das elites**. Rio de Janeiro: Fundação Getúlio Vargas, 2006.

HODGE, G. **Systems of knowledge organization for digital libraries**: beyond traditional authority files. Washington, D.C.: Clir Publication, 2000.

LANGRIDGE, D. **Classificação**: abordagem para estudantes de Biblioteconomia. Trad. Rosali P. Fernandez. Rio de Janeiro: Interciência, 1977.

LECOURT, D. (dir.). Classification. In: _____. **Dictionnaire d'histoire et philosophie des sciences**. Paris: Quadriga / PUF, 2006.

LESK, M. **Practical digital libraries: books, bytes, and bucks**. San Francisco: Morgan-Kaufmann, 1997.

MANN, M. Introduction to cataloguing and the classification of books. 2nd ed. Chicago: ALA, 1950.

MIRANDA, M. L. C. de; PARANHOS, J. P. B.; OLIVEIRA, J. X. de; PAES, M. S. A organização e representação do conhecimento em religiões yorubanas na Library of Congress Subject Heading. In: GUIMARÃES, J. A. C; DODEBEI, V. (Org.). **Desafios e perspectivas científicas para a organização e representação do conhecimento na atualidade**. [recurso eletrônico]. Marília: ISKO-Brasil: FUNDEPE, 2012. 285p. p. 153-159. e-Book.

PERROT, M. O inspetor Bentham. In: BENTHAM, Jeremy. **O panóptico**. Organização e tradução de Tomaz Tadeu a Silva. Belo Horizonte: Autêntica, 2000.

PIEIDADE, M. A. R. **Introdução à teoria da classificação**. 2. ed. Rio de Janeiro: Interciência, 1983.

RANGANATHAN, S. R. **Prolegomena to library classification**. Bombay: Asia Publ. House, 1967.

SAYERS, W. C. B. **A manual of classification for librarians and bibliographers**. 3rd ed. rev. with some corrections; with illustrations and bibliography. London: Grafton, 1962.

STONE, L. Prosopography. **Daedalus**, v.100, n.1, p.46-79, 1971.

THIESEN, I; MIRANDA, M. L.C. de. Jeremy Bentham, o utilitarismo e a classificação do conhecimento: elementos históricos para os estudos da Ciência da Informação. Rio de Janeiro, **Datagrama**, v.11, n.5, 2010. Disponível em: <http://www.dgz.org.br/out10/F_I_art.htm>.

THIESEN, I; MIRANDA, M.L.C. de. Jeremy Bentham e a classificação do conhecimento: elementos para as bases histórico-epistemológicas da Ciência da Informação. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, X, 2009, João Pessoa – PB. **Anais...** João Pessoa: UFPB, 2009.

UNITED STATES OF AMERICA. **Library of Congress Subject Headings**. 36th. ed. Washington, D.C., 2014. 6v.

MODELO DE COLABORAÇÃO PARA INDEXAÇÃO DE RECURSOS WEB

COLLABORATIVE MODEL FOR WEB RESOURCE INDEXING

Marcel Ferrante Silva
Gercina Angela Borém de Oliveira Lima

Resumo: O estudo apresenta um modelo de colaboração para a indexação de recursos *web*, oferecendo uma estratégia complementar à etiquetagem, através da seleção de termos de uma taxonomia facetada. Para a concepção do modelo de colaboração foi realizada uma pesquisa bibliográfica das seguintes temáticas: a) processo e modelos de recuperação da informação; b) indexação através de vocabulários controlados; c) classificação facetada; d) taxonomia facetada; e) indexação automática; f) navegação facetada; g) modelos de colaboração da *web* 2.0; h) sistemas colaborativos; i) fluxo de trabalho; j) moderação; e k) etiquetagem social. Na revisão de literatura foram pesquisados trabalhos sobre a etiquetagem facetada e classificação facetada colaborativa. Para a avaliação do modelo de colaboração foi realizado um teste de usabilidade. Apesar de uma grande parcela dos participantes do teste terem apontado sua preferência pela etiquetagem, em mais da metade das ações de indexação foi utilizada a taxonomia facetada, o que permitiu constatar que a utilização da mesma é válida para complementar a indexação.

Palavras-chave: classificação facetada, folksonomia, *web* 2.0, colaboração, indexação

Abstract: The study points to a collaborative model for indexing web resources, offering a complementary strategy labeling, through the selection of terms of a faceted taxonomy. For the design of the collaborative model a literature search was performed of the following themes: a) process models and information retrieval; b) indexing using controlled vocabularies; c) faceted classification; d) faceted taxonomy; e) automatic indexing; f) faceted navigation; g) collaboration templates of web 2.0; h) collaborative systems; i) workflow; j) moderation; and k) social labeling. In the literature review work on collaborative tagging and faceted faceted classification were investigated. For the evaluation of the collaborative model a usability test was performed. Although a large portion of the test participants have indicated their preference for labeling, in more than half of the shares indexing faceted taxonomy was used, which evidenced that the use of the same is valid to improve indexing.

Keywords: faceted classification, folksonomy, web 2.0, collaboration, indexing

1 INTRODUÇÃO

O ato de indexar é um processo complexo, pois, além de toda a subjetividade da análise prévia, em que são identificados os assuntos principais do documento (análise conceitual), o indexador precisa representá-lo através dos termos preferenciais contidos em um vocabulário controlado utilizado pelo sistema de recuperação da informação (SRI). Os sistemas de classificação e taxonomias são um tipo de vocabulário controlado, que possuem os termos organizados em uma estrutura hierárquica, geralmente representando a relação gênero-espécie ou todo-parte. Segundo Piedade (1983), esses sistemas têm origem nas classificações filosóficas, cujo precursor foi Aristóteles (384-322 a.C.). Na Biblioteconomia e Ciência da Informação (BCI), os sistemas de classificação enumerativos mais difundidos são:

a Classificação Decimal de Dewey (CDD) e a Classificação Decimal Universal (CDU). Esses sistemas foram criados no final do século XIX e se popularizaram durante o século XX e trouxeram a vantagem de agrupar hierarquicamente os documentos que apresentam temas relacionados, trazendo a possibilidade de mapear os termos de uma mesma área. Sendo assim, o usuário pode conhecer os termos mais gerais que são fornecidos inicialmente na raiz da hierarquia. Ao selecionar um termo, é possível acessar os termos inferiores, mais específicos que o termo superior, e prosseguir sucessivamente, até encontrar o termo com o assunto e especificidade desejados. Isso também permite a descoberta do conhecimento, em que o usuário, durante o processo de busca, toma conhecimento de termos que ele não conhecia anteriormente. Entretanto, Batista e Carvalho (2003, p. 40) apontam: “Uma estrutura hierárquica é fácil de compreender e usar, porém não é flexível”. Corroborando com essa afirmativa, Broughton e Slavic (2007) afirmam que a estrutura hierárquica nem sempre é flexível o suficiente para acomodar novos assuntos ou assuntos compostos.

Como alternativa a essa estrutura pouco flexível, surgiu um pouco mais tarde, a classificação facetada. O princípio da análise facetada foi idealizado por Ranganathan e propõe que os termos sejam organizados em facetas, ao invés de ter uma única grande hierarquia que acomoda todos os termos. Assim, uma faceta é uma lista de termos com as mesmas características, que servem para organizar objetos a partir de um determinado ângulo, perspectiva, ou dimensão (HEDDEN, 2010). Os assuntos compostos são gerados na indexação ou recuperação do documento, tornando as facetas mais sintéticas. Segundo Lima (2004, p. 78): “[...] em uma área do conhecimento, o número de assuntos específicos que podem ser selecionados numa classificação é infinito”.

De qualquer maneira, a indexação que utiliza um sistema de classificação enumerativo ou facetado geralmente é realizada manualmente por especialistas, acarretando um alto custo e lentidão do processamento humano, que necessita de treinamento e aprendizado do domínio para escolha dos termos mais adequados para função de representação e recuperação da informação (CARROLL; ROELOFFS, 1969 citado por VIEIRA, 1988). Paralelamente, a evolução exponencial da tecnologia computacional impulsionou a área de catalogação com o surgimento, no final da década de 1950, de uma nova técnica para a representação dos documentos: a indexação automática. Ao invés de ser realizada manualmente pelo indexador, a mesma é feita pelo computador através da linguagem natural. A indexação automática utiliza métodos que realizam a extração das palavras contidas no texto para formação do índice que, posteriormente, é utilizado na recuperação. Permite a busca por palavra-chave, que procura realizar um casamento entre as palavras digitadas pelo usuário e as palavras do

índice, que apontam para os documentos que as contêm. Tal representação tem maior exaustividade, com uma grande cobertura dos assuntos tratados no documento.

Na década de 90, viu-se a internet se popularizar após a invenção da *World Wide Web*. De repente, observou-se uma revolução significativa, em que uma pessoa poderia produzir e disponibilizar um documento para ser acessado em qualquer lugar do mundo de forma quase instantânea. Dias (1999) compara isso à invenção da imprensa. O número de sites aumentou em ritmo acelerado. No final da década de 90, dois projetos se destacaram como instrumentos para encontrar sites na internet: a gigante Yahoo!, que iniciou em 1994 com o site *Jerry and David's Guide to the World Wide Web*, sendo o catálogo de sites seu primeiro serviço oferecido; e o *Open Directory Project* (ODP) lançado pela Netscape. Ambos consistiam em uma grande hierarquia utilizada para indexar os sites e aceitavam sugestões fornecidas pelos usuários. No início dos anos 2000, esses catálogos ultrapassaram a marca de um milhão de sites catalogados. Entretanto, outro movimento estava prestes a acontecer.

Em 2001, ocorreu a explosão da bolha da internet com a queda de ações, fechamentos e fusões das empresas de tecnologia da informação baseadas na internet, as chamadas empresas "pontocom". Em 2004, a partir da percepção de características comuns às empresas que sobreviveram à bolha, teve origem o conceito *web 2.0*. Uma dessas características é apostar na colaboração dos usuários ou na inteligência coletiva que eles podem agregar.

Novamente, a evolução tecnológica impactou a área, permitindo que os próprios usuários, de maneira livre, sem regras, políticas ou vocabulários controlados, realizassem a indexação. O nome que se dá a esse tipo de indexação é folksonomia, caracterizada pela etiquetagem, que é a associação de etiquetas (termos na linguagem natural) a recursos *web*. É uma característica comum aos sites de maior sucesso da era *web 2.0*, tais como o Delicious, Digg, Wikipedia, Flickr, e Youtube. A recuperação da informação nesse caso baseia-se no casamento entre os termos usados na marcação do recurso *web* com os termos usados na busca pelo usuário. Um recurso *web* é qualquer item (página, imagem, vídeo, música) que tenha um *Uniform Resource Identifier* (URI). O URI pode ser uma *Uniform Resource Locator* (URL), que é um endereço de internet, ou um *Uniform Resource Name* (URN), que pode ser um endereço de *e-mail*. A etiquetagem pode ser considerada uma indexação social (GUEDES, 2010). Age de maneira similar à busca por palavra-chave e indexação automática, que também usa a linguagem natural. É adequada ao ambiente *web*, devido à "diminuição de custos e tempo para o usuário. Isto acontece pelo fato de não existirem hierarquias complexas de termos e nomenclaturas que o usuário tenha que aprender" (GUEDES, 2010, p.96).

Nessa época, estimuladas pelo desenvolvimento da internet e da *web* 2.0, também surgiram pesquisas relacionadas com a classificação facetada, e esse assunto ganhou uma nova roupagem no mundo digital. Com a aplicação da classificação facetada no meio digital surgiu a possibilidade de combinação dos termos de facetas diferentes para a formação dos assuntos compostos no ato de uma busca, o que se denominou de navegação facetada. Pode-se afirmar que a navegação facetada é uma estratégia eficiente de recuperação da informação e apresenta várias vantagens, conforme evidencia a revisão de literatura feita por Fagan (2010). A navegação facetada tornou-se bastante comum, principalmente em sites de comércio eletrônico, catálogos de endereços e bancos de imagens (DENTON, 2003).

Entretanto, a classificação facetada também pode apresentar obstáculos, conforme aponta Maly *et al.* (2010):

[...] para coleções que crescem em volume e variedade, o maior desafio envolvendo a classificação facetada é reclassificar os objetos quando ocorre uma modificação no sistema de classificação facetada. Um sistema de classificação com uma gestão centralizada, frequentemente apresenta dificuldades para ser adaptado em grandes coleções que crescem continuamente [...] (MALY *et al.*, 2010, p.70, tradução nossa).

A etiquetagem social apresenta-se como uma alternativa para tratar o problema acima, mas ela também tem limitações. Como exposto por Lai (2009), é difícil navegar em uma nuvem de etiquetas, formada por um conjunto superior a 100 (cem) etiquetas, pois apenas algumas etiquetas ficam em evidência, ofuscando as demais. Além disso, o autor expõe que as ambiguidades no uso das etiquetas podem remeter o usuário a informações pouco relevantes.

Nos anos 2000, surgiram pesquisas que apontam a aplicação da classificação facetada como forma de auxiliar a estruturação da etiquetagem e melhorar a recuperação de informação. Como exemplo desses estudos podem ser citados os protótipos Faccete (LAI, 2009) e o FaceTag (QUINTARELLI *et al.*, 2007) que permitem a etiquetagem facetada, orientando o usuário no momento de atribuir as etiquetas, dando a ele a possibilidade de escolher em qual faceta o termo será inserido. Em outro estudo, Spiteri (2011) aponta o emprego das facetas na etiquetagem, como uma possibilidade de melhorar a qualidade da indexação e recuperação da informação.

Maly *et al.* (2010), por sua vez propuseram um sistema de classificação facetada colaborativo, cuja vantagem é permitir que o sistema se reorganize automaticamente conforme os interesses dos usuários, ajudando-os a navegar pela coleção de forma rápida e intuitiva. Entretanto, a etiquetagem é uma abordagem para indexação já consolidada na *web*, e é usada por grandes *sites* da internet. Alterar ou trocar a etiquetagem por outra forma de indexação pode significar um risco para os administradores dos *sites* que já obtêm resultados

satisfatórios com a abordagem atual. Assim, uma opção é manter a etiquetagem e disponibilizar uma forma opcional de indexação. Um exemplo é o *site* Youtube, que permite a indexação tanto através da etiquetagem, quanto através de categorias já pré-determinadas. Outra alternativa, é manter a etiquetagem e permitir a indexação através da seleção de termos de uma taxonomia facetada. Essa abordagem pode trazer ganhos na indexação, e também permite a navegação facetada para a recuperação da informação.

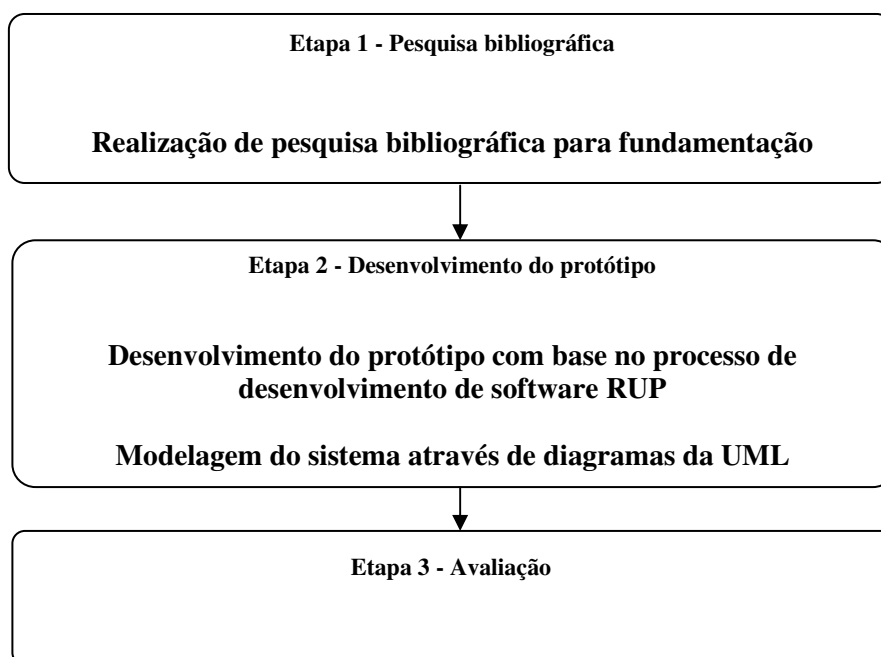
Partindo dos pontos levantados por estudos anteriores e questões apontadas em trabalhos de outros autores, a presente pesquisa teve como expectativa trazer respostas às seguintes questões: Como envolver colaborativamente os usuários na indexação desses recursos através de uma taxonomia facetada? A presença de uma taxonomia facetada é válida para os usuários na indexação de recursos *web*? Ao disponibilizar a taxonomia facetada e a etiquetagem para os usuários, qual é preferência na realização da indexação?

O objetivo desse estudo foi propor um modelo de colaboração para a indexação de recursos *web* que combine a etiquetagem e a seleção de termos de uma taxonomia facetada e verificar a preferência do usuário durante o processo de indexação.

2 METODOLOGIA

O percurso metodológico para atingir o objetivo da pesquisa é ilustrado na FIG. 1.

FIGURA 1 - Etapas da pesquisa



Fonte: Elaborado pelo autor.

Nas próximas seções, essas etapas são descritas detalhadamente.

2.1 Pesquisa bibliográfica

Primeiramente foram abordados os seguintes assuntos: processo e modelos de recuperação da informação (LANCASTER, 1993; BAEZA-YATES; RIBEIRO-NETO, 1999); indexação através de vocabulários controlados (NISO, 2005; NAVES, 2004); classificação facetada (CAMPOS, 2001; BATISTA; CARVALHO, 2003; LIMA, 2004; ARAUJO, 2005); taxonomia facetada (TZITZIKAS *et al.*, 2002; HEDDEN, 2010; MACULAN, 2011); indexação por meio da indexação automática (WIVES, 2000; BORGES, 2009); navegação facetada, com os tradicionais estudos da área: Flamenco (YEE *et al.*, 2003); Endeca (TUNKELANG, 2006); e Facetmap (SMITH *et al.*, 2006). Foram observados estudos recentes sobre navegação facetada em bases de dados RDF: FATIC (TVAROZEK; BIELIKOVÁ, 2007); HealthFinland (SUOMINEN *et al.*, 2007); BrowseRDF (OREN *et al.*, 2006); e Explorator (ARAÚJO, 2009, p.50) e catálogos *web* facetados: Niu (2012), Yang (2010); Tzitzikas *et al.* (2002); Makela *et al.* (2005); e FaThumb (KARLSON *et al.*, 2006).

A colaboração na *Web 2.0* foi chave para o sucesso de muitos sites em escala global, em especial aqueles em que o conteúdo é criado pelos usuários (VICKERY; WUNSCH-VICENT, 2007), e é um dos assuntos centrais nesse estudo. Está diretamente relacionada à etiquetagem social, na qual foram consultados os estudos de Catarino e Baptista (2007); Lopes (2002); Aquino (2007); e Guedes (2010). Para representar, classificar e analisar a colaboração da *Web 2.0* foram abordados estudos referentes a sistemas colaborativos (ELLIS *et al.*, 1991; DAMES, 2004), o modelo de colaboração 3C (LOTUS, 1995; GEROSA, 2006; PIMENTEL, 2006) e fluxo de trabalho (CASATI *et al.*, 1995; HOLLINGSWORTH, 1995; GEORGAKOPOULOS *et al.*, 1995). A partir desses estudos, buscou-se formular um modelo de colaboração para indexação de recursos em um catálogo *web*.

Também foram pesquisados estudos que fazem a união das de assuntos anteriores, principalmente a classificação facetada e a colaboração na *Web 2.0*, para compor a revisão de literatura. As seguintes pesquisas encontradas usavam a etiquetagem facetada: Faccete (LAI, 2009), FaceTag (QUINTARELLI *et al.*, 2007); e o *Faceted Index Internet Gateways* (DEVADASON *et al.*, 2002). Os estudos TagExplorer (SIGURBJÖRNSSON; ZWOL, 2010) e *Collaborative Faceted Categorization System* (MALY *et al.*, 2010, p.70) acrescentaram à etiquetagem social, a utilização da base de dados léxica *Wordnet* para o mapeamento de termos nas hierarquias das facetadas.

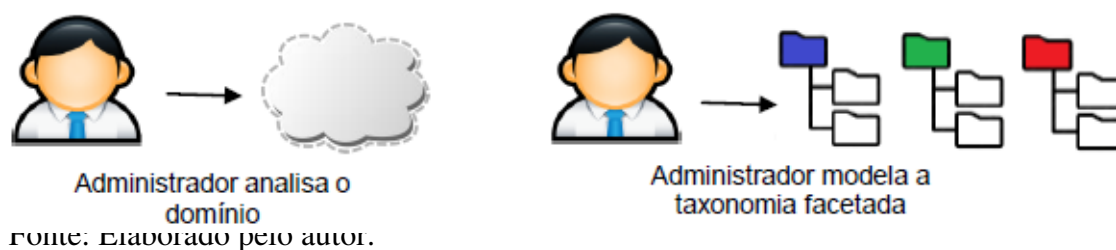
O presente estudo faz uso de vários elementos desses estudos e aponta outra abordagem, combinando a etiquetagem social com a taxonomia facetada para indexação dos recursos *web* por meio de um modelo de colaboração, que é detalhado nas próximas seções.

3 DESENVOLVIMENTO DO PROTÓTIPO

Foi adotado o processo de desenvolvimento de software *Rational Unified Process* (RUP) e a linguagem de modelagem *Unified Modeling Language* (UML) para a implementação do protótipo. O processo de desenvolvimento também foi inspirado por um estudo similar, que tinha o objetivo de criar um ambiente que permitisse a elaboração colaborativa de grandes documentos através de um sistema de gestão de conteúdo (SANTOS, 2009).

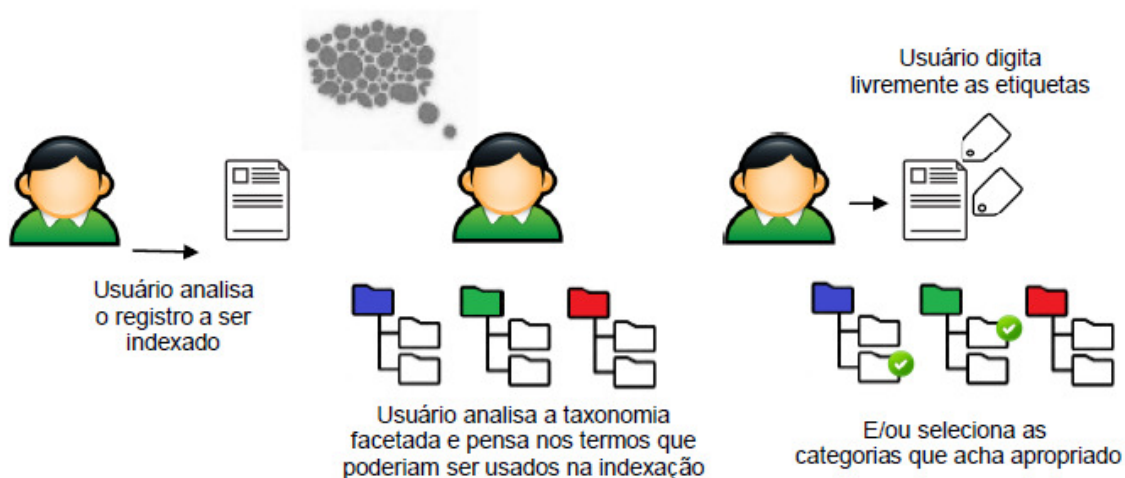
A seguir apresentamos o modelo de colaboração idealizado para indexação dos recursos. Previamente a esse processo, é necessário que um administrador determine as categorias e as facetas conforme o domínio de aplicação. É pressuposto que o administrador seja um especialista da área, com conhecimentos sobre os princípios da classificação facetada e indexação para a modelagem da taxonomia facetada (FIG. 2).

FIGURA 2 - Processo de Preparação



Na segunda etapa, o usuário pode realizar a indexação dos recursos usando a etiquetagem livremente ou a seleção de categorias da taxonomia facetada (FIG. 3). Os usuários atuam colaborativamente na indexação dos recursos, mas a modelagem da taxonomia facetada é restrita aos administradores.

FIGURA 3 - Processo de Indexação



3.1 MODELAGEM DE TAXONOMIA FACETADA

Foi escolhido para a avaliação do protótipo, aplicar o modelo de colaboração em um catálogo *web* facetado de empresas de uma cidade, fonte de informação também conhecida como páginas amarelas. Cada empresa foi modelada como um recurso *web*, com uma URL própria. Para a realização da classificação facetada foram usados os atributos das empresas. Foi gerada uma lista de empresas de Goiânia e conforme os exemplos de catálogos *web* facetados de empresas da fundamentação teórica, foram determinadas as seguintes facetas para a taxonomia facetada: área de atividade e bairro. Pode-se entender que essas facetas correspondem respectivamente às categorias fundamentais Personalidade e Espaço do PMEST de Ranganathan. A área de atividade tinha 46 subcategorias e dois níveis hierárquicos e bairro tinha 356 subcategorias em um único nível hierárquico. A modelagem ficou limitada nessas duas facetas devido aos atributos disponíveis das empresas na listagem que foi obtida, entretanto, essas são as facetas mais comuns em catálogos de empresas na internet. Outro ponto importante é que o objetivo do estudo não é validar a taxonomia facetada e sim a preferência do usuário no seu uso durante o processo de indexação.

3.2 PROCESSO DE INDEXAÇÃO DO RECURSO

A partir da modelagem, elaboração de requisitos e diagramas, foi desenvolvido um protótipo da proposta do estudo. É possível visualizar a interface desse protótipo na FIG. 4. Após se registrar e fornecer seu usuário e senha, o usuário é redirecionado para indexar um recurso, evento que consiste em: inserir as etiquetas (palavras na linguagem natural do usuário) e/ou selecionar as categorias da taxonomia facetada que o recurso pertence.

FIGURA 4 - Interface de indexação do recurso

Palavra-chaves (aperte ENTER ou Espaço para inserir uma palavra-chave)

Etiquetagem Social

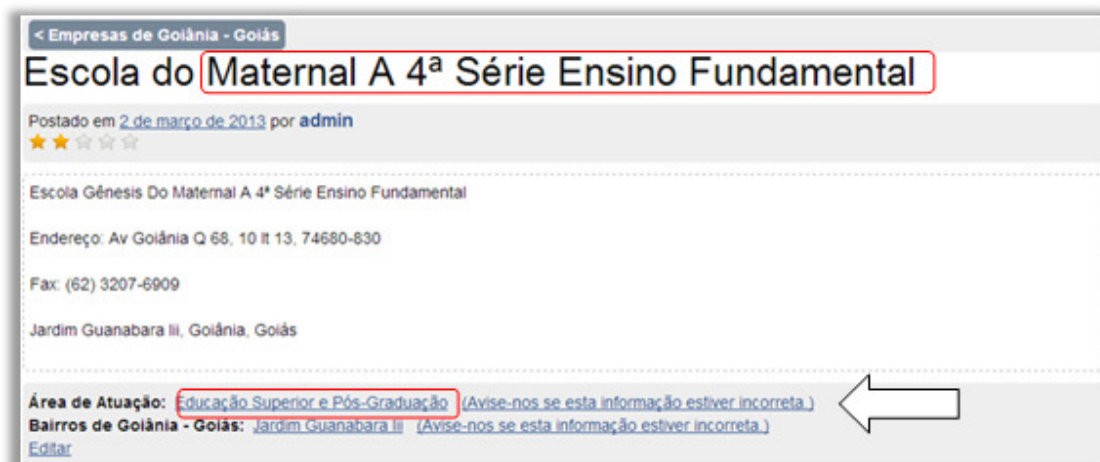
Classifique sua empresa segundo os critérios abaixo

Fonte: Elaborado pelo autor.

3.3 PROCESSO DE REVISÃO DA INDEXAÇÃO DO RECURSO

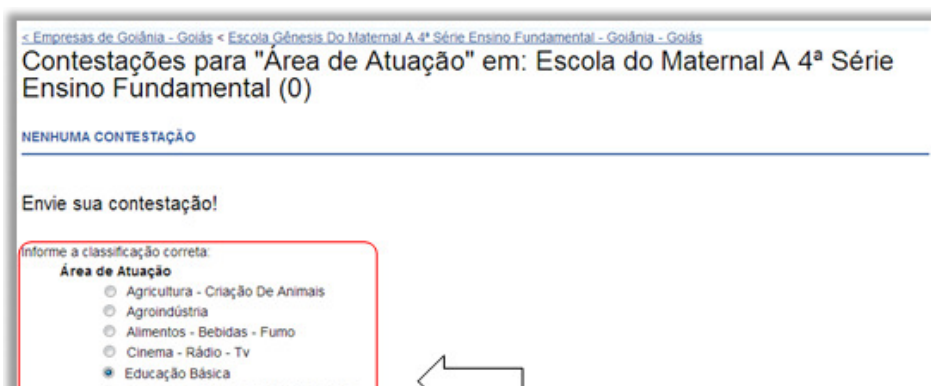
O usuário também pode revisar a indexação realizada em um recurso cadastrado por outro usuário. Esta ação foi denominada de contestação, no sentido de que o usuário reporta que um recurso está com uma indexação incorreta ou incompleta clicando em "Avisenos se esta informação estiver incorreta" (FIG. 5).

FIGURA 5 - Visualização de um recurso



Clicando nessa opção, o usuário é direcionado para a página de contestação, na qual ele informa a categoria correta do item em questão (FIG. 6).

FIGURA 6 - Contestação de classificação de um recurso



Fonte: Elaborado pelo autor.

Nesse caso, a indexação realizada é submetida à aprovação por um administrador (moderação). Na área de administração, que somente os administradores têm acesso, podem ser observadas as contestações informadas pelos usuários. A interface exibe, por padrão, as contestações que ainda não foram tratadas (FIG. 7).

FIGURA 7 - Tratamento das contestações

Contestações			
Não tratadas (4) Tratadas (7)			
Idéia	Contestação	Galeria	Situação
Escola Maternal A 4ª Série Ensino Fundamental	A escola tem apenas ensino fundamental. [...]	Empresas de Goiânia - Goiás	Não Verificada
Básico Materiais Construção	Área da Construção civil [...]	Empresas de Goiânia - Goiás	Não Verificada
Condomínio Do Edifício West Plaza	imóveis [...]	Empresas de Goiânia - Goiás	Não Verificada
Condomínio Do Edifício West Plaza	setor bueno [...]	Empresas de Goiânia - Goiás	Não Verificada

Clicando na contestação, o administrador aprova ou rejeita a contestação sugerida pelo usuário. A contestação pode servir para adicionar ou alterar uma classificação de um recurso (FIG. 8). Se aprovado, a classificação do recurso é alterada para todos os usuários.

FIGURA 8 - Analisando contestação

Verificar Contestação

Verificar Contestação de Idéias

Conteúdo da contestação

Classificação contestada: > Área de Atuação > Educação Superior e Pós-Graduação
 Classificação sugerida: > Área de Atuação > Educação Básica
 Descrição:
 A escola tem apenas ensino fundamental.

A contestação está confirmada por uma fonte confiável?

Sim Não

Observação:

OK

Fonte: Elaborado pelo autor.

4 AVALIAÇÃO E RESULTADOS

4.1 Teste de usabilidade

Dentre as técnicas para avaliação foi escolhido o teste de usabilidade, em que são observadas as ações dos usuários na realização de tarefas, por duas razões principais: busca-se conhecer o comportamento do usuário na realização da tarefa de indexação em uma interface que oferece duas estratégias opcionais, a etiquetagem e a taxonomia facetada e também na revisão da indexação de um recurso de outro usuário. O ambiente de teste escolhido foi o laboratório de informática da Faculdade de Informação e Comunicação (FIC) da Universidade Federal de Goiás (UFG), devido à facilidade de acesso aos participantes.

Trata-se de um teste de usabilidade que utiliza o método de observação indireta, por ser menos invasivo, no qual as ações dos usuários são gravadas através da captura do vídeo do computador. Para a realização do teste cogitou-se utilizar o *software* Morae, da Camtasia, que é dedicado a testes de usabilidade. Entretanto, era necessário gravar as ações dos usuários nos computadores do laboratório de informática da FIC, cujos computadores têm o sistema operacional Linux, com a distribuição Ubuntu. Assim, após reunião com a equipe de informática da FIC, foi escolhido o programa *RecordMyDesktop* que funciona nesse sistema operacional e que faz a captura das telas do computador com suporte a vários formatos de vídeo. A instalação do *RecordMyDesktop* foi muito simples, com a busca do programa pela própria ferramenta de pesquisa de aplicativos do Ubuntu. Após a instalação, aparece um ícone no canto esquerdo inferior da tela, que dá acesso ao aplicativo.

Para realização das tarefas, o protótipo desenvolvido foi acessado através do navegador de Mozilla Firefox, disponível na distribuição do Ubuntu. Foi elaborado um termo de consentimento entregue para os participantes antes do início da sessão de testes para a autorização do uso das informações coletadas durante os mesmos. Optou-se em utilizar alunos do curso superior, como participantes do teste, com o intuito de trabalhar com um perfil mais uniforme e com pessoas com pré-requisitos mínimos. Foram selecionados dez alunos voluntários do quarto período do curso de graduação em Gestão da Informação da UFG, sendo que oito participaram efetivamente do teste. Esse número de participantes estava de acordo com Nilsen (2000) que recomenda o mínimo de cinco participantes para realização do teste de usabilidade. O fato de serem estudantes de graduação permitiu que os usuários

escolhidos tivessem um conhecimento básico, com experiência de, no mínimo, dois anos em navegação pela internet.

4.1.1 Lista de tarefas

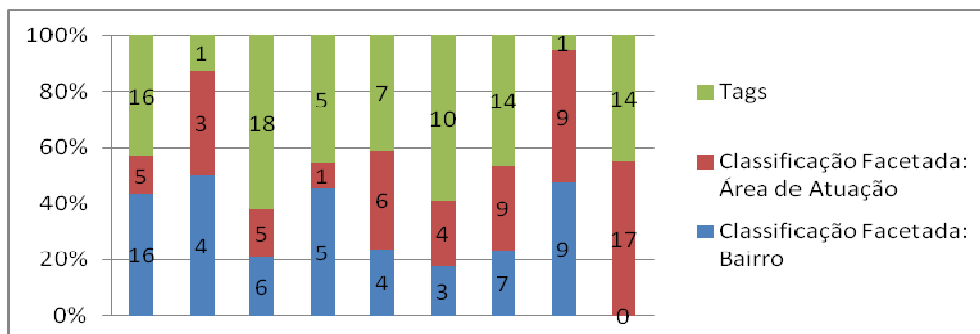
A lista de tarefas planejada consistia das seguintes tarefas: 1) cadastrar empresas a partir de uma lista de empresas fornecidas. Não foi fornecida a área de atividade da empresa, que deveria ser deduzida a partir do seu nome completo. Os outros dados fornecidos sobre a empresa foram: telefone, endereço, bairro, CEP e cidade. Foi fornecida uma lista com cerca de vinte empresas de regiões com áreas de atividade aleatórias. Foram dados 10 minutos para o usuário cadastrar a quantidade de empresas que conseguisse e foi explicitamente mencionado que o usuário poderia utilizar as categorias fornecidas ou usar suas próprias palavras (palavras-chave) para descrevê-las; 2) revisar as classificações que um outro usuário realizou. Também foi estipulado 10 minutos para realização dessa atividade. O usuário revisor poderia apontar qualquer erro na classificação que fosse detectado, sendo que, um terço das empresas já previamente cadastradas haviam sido, propositalmente, indexadas de maneira indevida.

4.2 Análise da indexação dos recursos

Na primeira atividade de cadastro e indexação manual de empresas, foi avisado para todos os usuários que participaram do teste, que a indexação poderia ser feita da maneira como preferisse, utilizando etiquetas ou selecionando categorias das facetadas. Existiam duas facetadas: uma que se referia à área de atividade, e outra referente ao bairro em que se localizava a empresa.

A primeira observação com relação à utilização de termos da taxonomia facetada (classificação facetada) versus a utilização de etiquetas livremente digitadas pelos usuários é que, para o cadastro do bairro, os usuários utilizaram as categorias em todos os casos. Uma possível explicação para isso é o fato de o bairro ser uma informação que apresentava correspondência absoluta com uma das categorias da faceta “bairros”. Com relação à área de atuação, o usuário deveria procurar na taxonomia da faceta qual categoria era mais apropriada para tal recurso, ou poderia digitar etiquetas, usando a linguagem natural. O GRÁF. 1 evidencia que a maioria dos usuários preferiu utilizar as etiquetas a usar a classificação facetada para registrar esse tipo de informação.

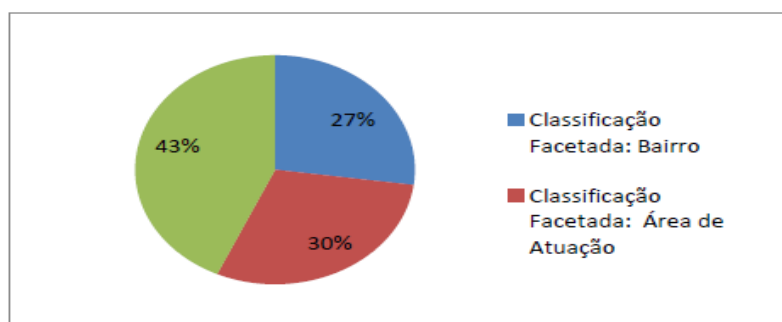
GRÁFICO 1 – Utilização da classificação facetada versus etiquetas por usuário



Fonte: Elaborado pelo autor.

A indexação das áreas de atuação foi realizada por meio de etiquetas em 43% dos casos, contra 30% em que houve uma seleção de categorias da faceta. Não houve nenhum usuário que usou etiquetas para informar o bairro da empresa (GRÁF. 2).

GRÁFICO 2 – Total da utilização da classificação facetada versus etiquetas

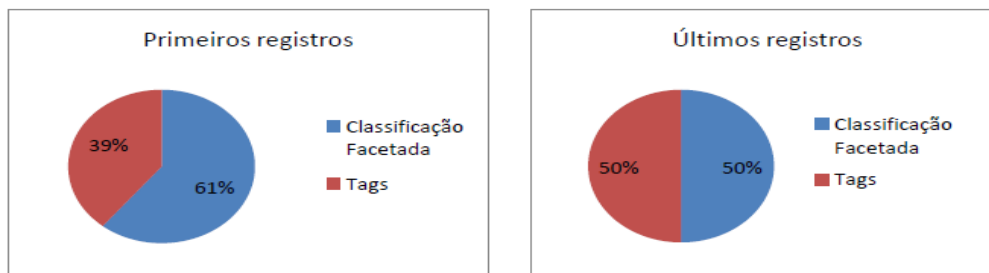


Fonte: Elaborado pelo autor.

4.2.1 Alteração no comportamento do usuário

Foi percebida uma mudança no comportamento do usuário durante a realização da indexação. Observou-se que, nos últimos recursos, houve grande preferência pela utilização das etiquetas. Nesses, a maioria dos usuários atribui um termo do nome próprio da empresa na sua etiquetagem (por exemplo: depósito, imobiliária ou madeireira). Isso pareceu ser mais simples que escolher e selecionar uma categoria na faceta (GRÁF. 3) durante a indexação do recurso. Outra consideração que merece destaque é que 44% dos usuários procuraram categorias na taxonomia das áreas de atuação, mas não encontraram nenhuma adequada para a indexação. Isso pode ser devido à ausência da categoria na taxonomia, falta de atenção ou pressa na execução da tarefa. Nessa situação, o usuário optou geralmente em criar etiquetas.

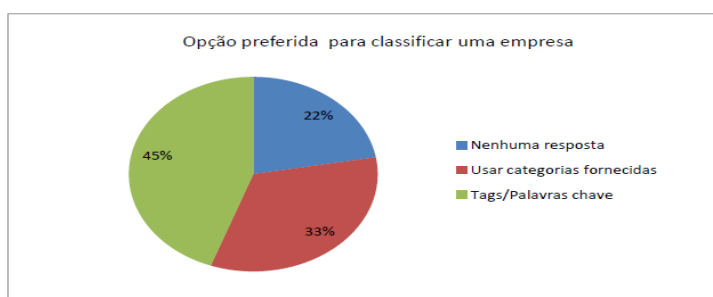
GRÁFICO 3 – Utilização das categorias versus etiquetas



4.2.2 Questionário de satisfação

A partir do questionário de satisfação do usuário, foi elaborado o GRÁF. 4 que revelou que a preferência da maioria dos usuários era digitar as etiquetas, ao invés de usar as categorias já existentes para indexação. Um dos motivos disso pode ser constatado durante a observação dos vídeos, e em parte se deve ao fato de que as facetas eram demasiadamente grandes e a interface não dispunha de nenhuma ferramenta para filtrar ou localizar as categorias.

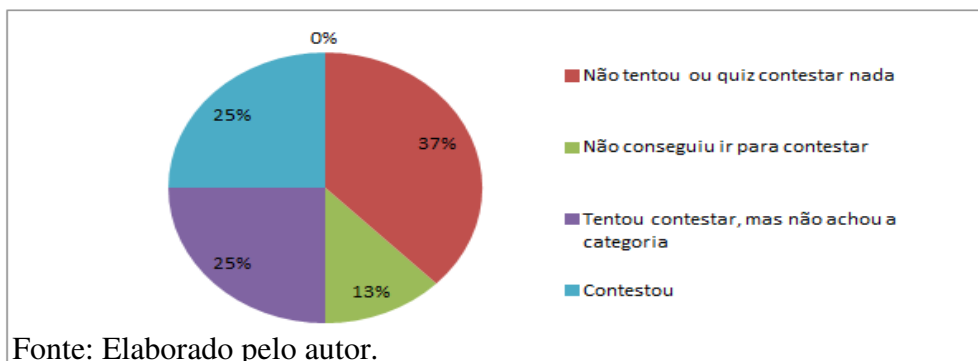
GRÁFICO 4 – Preferência: categorias facetadas versus etiquetas



4.3 Análise da revisão da indexação dos recursos

A tarefa de revisar e apontar um erro na classificação de um recurso ou adicionar uma nova classificação podia ser realizada por meio do *link* "Avise-nos se esta informação estiver incorreta ou incompleta". No GRÁF. 5, vê-se uma estatística referente ao comportamento dos usuários diante da solicitação de realizar essa revisão. Menos de um terço dos usuários (25%) conseguiu realizar essa ação. Atribui-se esse baixo número à dificuldade que tiveram em saber como realizar a ação (baixa usabilidade da interface), bem como identificar uma classificação errada ou mais apropriada.

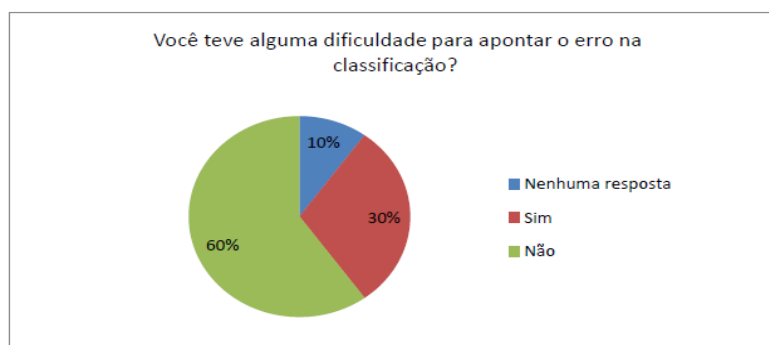
GRÁFICO 5 – Usuários revisam a classificação dos recursos



4.3.1 Questionário de satisfação

No GRÁF. 6, percebe-se que cerca de um terço dos usuários afirmaram ter tido dificuldade ao realizar a contestação (pedido de revisão da classificação de um recurso). Tal fato confirma o comportamento dos usuários observado na captura do vídeo de suas ações e indica a necessidade de melhorias na interface para a realização dessa ação com maior eficiência.

GRÁFICO 6 – Contestação



Fonte: Elaborado pelo autor.

5 CONSIDERAÇÕES FINAIS

O estudo propôs um modelo de colaboração para a indexação de recursos *web* que combina a etiquetagem e a seleção de termos de uma taxonomia facetada implementado através de um protótipo de um catálogo *web* facetado de empresas. Através do teste de usabilidade foi possível constatar que a taxonomia facetada é uma estratégia válida para indexação de recursos, mas não deve ser a única forma de indexação, pois a etiquetagem segundo o questionário de satisfação foi preferida por 45% dos usuários e segundo a captura do vídeo foi utilizada em 43% das ações de indexação. De qualquer maneira, isso significa que em soluções de larga escala, é possível ter um batalhão de usuários realizando a classificação facetada gratuitamente, o que se traduz em economizar com um processo

subjetivo e dispendioso quando usado o modelo tradicional através de especialistas. Conclui-se que a classificação facetada e a etiquetagem são estratégias complementares e é viável disponibilizar ambas na interface de indexação.

É importante ressaltar que essas estatísticas geradas foram para um caso particular, que combina um grupo de usuários com um perfil definido, um domínio específico, uma taxonomia facetada própria para esse domínio, tarefas criadas para realização desse experimento e uma interface com um design singular para esse estudo. A alteração de uma dessas variáveis pode alterar os resultados obtidos. Na revisão da indexação dos recursos de outros usuários, apesar da maioria dos usuários (63%) demonstrarem essa intenção, apenas 25% conseguiram efetuar tal tarefa. Dessa forma, em futuros trabalhos planeja-se melhorar a interface do protótipo com o objetivo que a tarefa de indexação seja desempenhada com o menor esforço possível e de maneira mais intuitiva. É previsto a realização de novos testes para avaliar se os resultados são repetidos para grupos de usuário com perfil heterogêneo, em maior escala, em outros domínios como o de bibliotecas digitais e arquivos abertos com diferentes taxonomias facetadas. Tem-se a intenção de avaliar se a taxonomia facetada melhora a qualidade da indexação, como a sua presença influencia na etiquetagem, se o seu tamanho, composição e domínio impactam no processo, em quais situações em que a etiquetagem é preferida em relação a taxonomia facetada, e os ganhos que esse modelo pode proporcionar na recuperação da informação.

Também é planejada a expansão do modelo de colaboração para a gestão da taxonomia facetada com a possibilidade do usuário sugerir novas categorias, correção ou exclusão de categorias já existentes.

REFERÊNCIAS

AQUINO, M. C. Hipertexto 2.0, Folksonomia e Memória Coletiva: Um estudo das tags na organização da *web*. **Revista da Associação Nacional dos Programas de Pós-Graduação em Comunicação**, p. 1-18, ago, 2007.

ARAÚJO, S. F. C. de. **Explorator**: ferramenta para exploração de dados RDF baseado em uma interface de manipulação direta. 2009. Dissertação (Mestrado em Informática) - Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2009.

BAEZA-YATES, R; RIBEIRO-NETO, B. **Modern Information Retrieval**. New York: Addison-Wesley, 1999.

BATISTA, E. A. **Uma Taxonomia Facetada para Técnicas de Elicitação de Requisitos**. 2003. Dissertação (Mestrado em Ciência da Informação) - Instituto de Comunicação, Unicamp, Campinas, 2003.

BLATTMANN, U.; SILVA, F. C. C. Colaboração e interação na *web 2.0* e biblioteca 2.0. **Revista ACB: Biblioteconomia em Santa Catarina**, v. 12, n. 2, p. 191-215, jul./dez., 2007.

BORGES, G. S. B. **Indexação automática de documentos textuais: proposta de critérios essenciais**. 2009. Dissertação (Mestrado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2009.

BROUGHTON, V.; SLAVIC, A. Building a faceted classification for the humanities: principles and procedures. **Journal of documentation**, 2007.

CAMPOS, M. L. de A. **A organização de unidades do conhecimento em hiperdocumentos: o modelo conceitual como um espaço comunicacional para realização da autoria**. 2001. Tese (Doutorado em Ciência da Informação) - Instituto Brasileiro de Informação em Ciência e Tecnologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2001.

CASATI, F. et al. Conceptual Modeling of WorkFlows. In: OOER'95 - 14th International Object-Oriented and Entity-Relationship Modelling Conference, 1995. **Proceedings...** Springer-Verlag, p. 341-354, 1995.

CATARINO, M. E.; BAPTISTA, A. A. Folksonomia: um novo conceito para a organização dos recursos digitais na Web. **DataGramZero - Revista de Ciência da Informação**, Rio de Janeiro, v. 8, n. 3, jun., 2007.

DAMES, K. M. **Social software in the library**. LLRX, jul. 2004. Disponível em: <<http://www.llrx.com/features/socialsoftware.htm>>. Acesso em: 15 dez. 2007.

DENTON, W. **How to make a faceted classification and put it on the web**. Toronto: Miskatonic University Press, 2003. Disponível em: <<http://www.miskatonic.org/library/facet-web-howto.html>>. Acesso em: 30 ago. 2010.

DEVADASON, F. J. et al. Faceted Indexing Based System for Organizing and Accessing Internet. **Knowledge Organization**, v. 29, p. 65-77, 2002.

DIAS, C. A. Hipertexto: evolução histórica e efeitos sociais. **Ciências da informação**, v. 28, n. 3, p. 267, set./dez., 1999.

ELLIS, C.A. et al. Groupware - Some Issues and Experiences. **COMMUNICATIONS OF THE ACM**, v. 34, n. 1, p. 38-58, 1991.

FAGAN, J. C. Usability studies of faceted browsing: A literature review. **Information Technology and Libraries**, v. 29, p. 58-66, 2010.

GEORGAKOPOULOS, D. et al. An Overview of Workflow Management: From Process Modeling to Workflow Automation Infrastructure. **Distributed and Parallel Databases**, Boston, v.3, p. 119-153, 1995.

GUEDES, R. M. **A abordagem dialógica na indexação social**. 2010. Dissertação (Mestrado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2010.

HEDDEN, H. **The Accidental Taxonomist**. Medford, NJ: Information Today, 2010.

HOLLINGSWORTH, D. **Workflow Management Coalition: The Workflow Reference Model**. Hampshire: Workflow Management Coalition, 1995.

KARLSON, A. K. et al. FaThumb: A Facet-based Interface for Mobile Search. **Navigation: Mobiles & Audio**, v. 22, n. 27, p. 711-720, apr., 2006.

LAI, P. **Facette: Using Facets to Improve Tag-based Bookmarking**. 2009. Dissertação (Master of Engineering in Electrical Engineering and Computer Science) - Massachusetts Institute of Technology, 2009.

LANCASTER, F. W. **Indexação e resumos: teoria e prática**. Brasília: Briquet de Lemos, 1993.

LIMA, G. A. B. O. **Mapa hipertextual (MHTX): um modelo para organização hipertextual de documentos**. 2004. Tese (Doutorado em Ciência da Informação) - Escola de Ciência de Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2004.

LOPES, I. L. Uso das linguagens controlada e natural em bases de dados: Revisão da literatura. **Ciência da Informação**, Brasília, v. 31, n. 1, p. 41-52, jan./abr., 2002.

LOTUS DEVELOPMENT CORPORATION. **Groupware: Communication, Collaboration and Coordination**. Cambridge: Lotus Development Corporation, 1995.

MAKELA, E. et al. Semantic Yellow Page Service Discovery: The Veturi Portal. In: 4th International Semantic Web Conference, 2005. **Proceedings...** Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.214.5211&rep=rep1&type=pdf>>. Acesso em: 02 set. 2011.

MALY, K. et al. A collaborative faceted categorization system - user interactions. In: 14th International Conference on Electronic , 2010, Helsinki. **Proceedings...** Helsinki: Transforming the Nature of Communication, 2010.

NAVES, M. M. L. **Curso de indexação: princípios e técnicas de indexação, com vistas à recuperação da informação**. Belo Horizonte: UFMG. 2004. Apostila.

NIELSEN, J. **Why You Only Need to Test with 5 Users**. Useit.com. 2000. Disponível em: <<http://www.useit.com/alertbox/20000319.html>>. Acesso em: 07 fev. 2007.

NISO-NATIONAL INFORMATION STANDARDS ORGANIZATION. Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies. Z39.19. Maryland: ANSI/NISO, 2005.

NIU, X. **Beyond text queries and ranked lists: faceted search in library catalogs**. 2012. Tese (Doctor of Philosophy) - School of Information and Library Science, University of North Carolina at Chapel Hill, 2012.

OREN, E. et al. **Extending faceted navigation for RDF data**. In: 5th International Semantic Web Conference, 2006. **Proceedings...** ISWC, 2006.

PIEDADE, M. A. R. **Introdução à teoria da classificação**. 2. ed., Rio de Janeiro: Interciência, 1983.

- QUINTARELLI, E. et al. FaceTag: Integrating Bottom-up and Top-down Classification in a Social Tagging System. **Bulletin of the American Society for Information Science and Technology**, p. 10-15, jun./jul., 2007.
- SANTOS, E. M. dos. **Um Modelo para Desenvolver Editores Colaborativos Hierárquicos**. 2009. Dissertação (Mestrado em Informática) - Instituto Tecnológico de Aeronáutica, São Jose dos Campos, 2009.
- SIGURBJÖRNSSON, B., ZWOL, R. **Tagexplorer**: Faceted browsing of flickr photos. Yahoo! Research, 2010. Technical Report.
- SMITH, G. et al. FacetMap: A Scalable Search and Browse Visualization. **IEEE Transactions on Visualization and Computer Graphics**, v. 12, n. 5, p. 797-804, 2006.
- SPITERI, L. **Faceted navigation of social tagging applications**. Dalhousie Repository. Halifax: Dalhousie University, 2011.
- SUOMINEN, O. et al. User-centric Faceted Search for Semantic Portals. **The Semantic Web: Research and Applications**. Springer Berlin Heidelberg, p. 356-370, 2007.
- TUNKELANG, D. Dynamic category sets: An approach for faceted search. In: ACM SIGIR Workshop On Faceted Search , 2006. **Proceedings...** Seattle, WA, 2006.
- TVAROZEK, M.; BIELIKOVAR, M. Personalized faceted navigation for collections. In: Semantic Media Adaptation and Personalization, Second International Workshop, 2007. **Proceedings...** IEEE, 2007.
- TZITZIKAS, Y. et al. Extended Faceted Taxonomies for Web Catalogs. In: Web Information Systems Engineering - WISE, Third International Conference, 2002. **Proceedings...** IEEE, 2002.
- VICKERY, G.; WUNSCH-VINCENT, S. **Participative Web And User-Created Content: Web 2.0 Wikis and Social Networking**. Paris: Organization for Economic Cooperation and Development - OECD, 2007.
- VIEIRA, S. B. INDEXAÇÃO AUTOMÁTICA E MANUAL: REVISÃO DE LITERATURA. **Ciência da Informação**, Brasília, v. 17, n. 1, p. 43-57, 1988.
- WIVES, L. **Tecnologias de Descoberta de Conhecimento em Textos aplicadas à Inteligência Competitiva**. 2002. Exame de Qualificação - Pós-Graduação em Computação, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2002.
- YANG, S. Q.; HOFMANN, M. A. The next generation library catalog: A comparative study of the OPACs of Koha, Evergreen, and Voyager. **Information Technology and Libraries**, v. 29, n. 3, p. 141-150, 2013.
- YEE, K. P. et al. Faceted metadata for image search and browsing. In: SIGCHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, 2003. **Proceedings...** ACM, 2003.

ABORDAGEM TEÓRICO-METODOLÓGICA NA ORGANIZAÇÃO DE IMAGENS EM PATOLOGIA

METHODOLOGICAL APPROACH TO ORGANIZATION OF IMAGES IN PATHOLOGY

Joice Cleide Cardoso Ennes de Souza
Elan Cardozo Paes de Almeida

Resumo: O objetivo da pesquisa é a investigação dos princípios de organização de imagens biomédicas de cortes histopatológicos com fins de armazenamento e recuperação em bancos de imagens com objetivos acadêmico-científicos, de modo a possibilitar futuras aplicações profissionais e educacionais para a área de Patologia. Procurou-se refletir sobre questões referentes à organização e representação de imagens de lâminas histopatológicas e sua recuperação e utilização como recurso didático e material de pesquisas futuras a partir de referencial teórico interdisciplinar, fundamentado na Ciência da Informação, na Teoria da imagem, na recuperação da imagem no âmbito da biomedicina e nas Teorias da Classificação Facetada e do Conceito (fundamentos teóricos para a representação da informação e do conhecimento). O critério de organização das imagens de lâminas histopatológicas pelas aulas ministradas apresenta um aspecto positivo para o discente, contudo, a posterior recuperação das imagens é prejudicada por estar associada a conteúdos programáticos específicos da disciplina de Patologia Geral, o que dificulta a busca futura por lesão ou tipo de agressão, por exemplo.

Palavras-chave: Imagens biomédicas. Organização da informação. Patologia Geral. Lâminas histopatológicas.

Abstract: The objective of the study is research the principles of organizing biomedical pictures of histopathological sections for the purpose of storage and retrieval in image databanks with academic and scientific goals, to enable future professional and educational applications to the area of pathology. Sought to reflect on issues relating to the organization and representation of images of histological slides and recovery and use as a teaching resource and future research material from interdisciplinary theoretical framework, grounded in Information Science, Theory of the image, image retrieval in the context of biomedicine and Theory of Faceted Classification and Theory of Concept (theoretical for the representation of information and knowledge bases). The arrangement of their images of histopathological slides for classes has a positive aspect for the student; however, the later retrieval of images is hindered by being associated with specific syllabus content of General Pathology, which complicates the search for future injury or type of aggression, for example.

Keywords: Biomedical images. Organization of information. General Pathology. Histopathological slides.

1 INTRODUÇÃO

No início do século XXI, a imagem fotográfica mantém sua primazia como recurso informacional. Desde sua criação, a fotografia se caracterizou por diferentes técnicas e suportes, culminando com a fotografia digital. Usada com fins históricos, educativos e científicos, a imagem fotográfica é uma fonte de informação para diferentes tipos de usuários e para diferentes finalidades.

O uso de recursos visuais na formação biomédica é prática acadêmica constante nos currículos dos cursos de graduação, sendo que nas últimas décadas há a predominância da utilização de imagens nas aulas expositivas e laboratoriais, adotadas tanto para a exposição do conteúdo como para contribuir com a obtenção de diagnósticos e definição da terapêutica a ser adotada no paciente, dependendo do caso estudado.

Tomada como forma de registro, no ambiente de ensino a imagem permitiu que professores e alunos tivessem acesso a informações não transmitidas em forma textual, uma vez que a decodificação do visual para o textual consiste em um desafio. Segundo Pinto (2000, p.42), "a imagem é uma técnica que permite a permanência temporal do registro de algo, mesmo quando essa coisa não se encontra mais presente", apoiando, dessa forma, o ensino do conteúdo biomédico com a exibição de imagens para ilustrar o tema da aula expositiva ou laboratorial.

A representação da imagem biomédica suscita os seguintes questionamentos: como organizar o acervo de imagens biomédicas, especificamente imagens digitalizadas de lâminas histopatológicas? O que devemos destacar na imagem? Como garantir a recuperação da imagem pelo usuário do sistema?

Para tanto, objetiva-se investigar princípios de organização de imagens biomédicas de cortes histopatológicos com fins de armazenamento e recuperação em bancos de imagens com objetivos acadêmico-científicos, a fim de possibilitar futuras aplicações profissionais e educacionais na área de Patologia. O escopo da pesquisa destaca o papel do profissional da informação como um especialista no tratamento da informação e como sua participação é fundamental para as pesquisas de ponta na área biomédica.

Iniciamos o artigo apresentando os aspectos relativos à organização e recuperação de imagens biomédicas nos sistemas de informação em saúde na *web*, para na seção seguinte analisar como a imagem é utilizada pelos profissionais da área da Patologia Geral na definição de diagnósticos. Na quarta seção apontamos conceitos da Ciência da Informação aplicados à organização de imagens biomédicas e fluxos informacionais da imagem em sistemas de informação. Os aspectos conceituais da análise da imagem são observados na seção dedicada à Teoria da Imagem. A abordagem metodológica para a organização das imagens de lâminas histopatológicas encontra-se na seção seis, onde focamos as teorias da Classificação Facetada e do Conceito na identificação de categorias.

2 ORGANIZAÇÃO E RECUPERAÇÃO DE IMAGENS BIOMÉDICAS NA WEB

Os desafios da organização e recuperação da imagem biomédica na *web* é objeto de estudo da informática biomédica, apresentada como um campo interdisciplinar que estuda e desenvolve dados, informações e conhecimentos biomédicos para a pesquisa científica e definição de diagnósticos (HSU *et al.*, 2013, p. 1010) em sistemas de informação em saúde.

Com a disseminação do uso da imagem como recurso da prática médica, a informática biomédica apresenta um subcampo em rápido crescimento denominado *biomedical imaging informatics* (BMII), informática de imagens biomédicas, responsável pelo processamento de imagens biomédicas, otimização e uso de imagens e sua aplicação em pesquisas clínicas diversas.

A informática de imagens biomédicas abrange todo tipo de imagem médica em qualquer área da medicina clínica, como radiologia, patologia, dermatologia, oftalmologia, entre outras. Sua rápida expansão deve-se à ampliação do uso de imagens em pesquisas científicas nas áreas de patologia, anatomia e genética. O uso da imagem como registro clínico está diretamente relacionado com o aumento de alternativas que permitem a identificação do diagnóstico e a personalização no cuidado do paciente. Contudo, o grande número de imagens biomédicas dificulta a recuperação da informação pelos pesquisadores e profissionais da área e, conseqüentemente, retardando a definição do diagnóstico e tratamento do paciente. A BMII tem como finalidade estudar formas de recuperação da informação nos bancos de imagens biomédicas, a partir da "interpretação da imagem" (AMERICAN MEDICAL INFORMATICS ASSOCIATION, 2014).

Hsu *et al.*(2013) esclarecem que o desenvolvimento da informática de imagens biomédicas exige um contato interdisciplinar com as ciências humanas e sociais para a avaliação de propostas e soluções técnicas. Tal afirmativa vai ao encontro das atividades atualmente desenvolvidas no âmbito da BMII: aquisição de imagem e seu gerenciamento, controle de qualidade, extração de informações (análise da imagem), modelagem, detecção e diagnóstico auxiliado por computador, recuperação de imagens baseado em conteúdo (CBIR) e visualização, pressupondo o entrosamento de profissionais de diferentes áreas.

A AMIA (*American Medical Informatics Association*) defende que a análise das imagens seja realizada pelas características de nível baixo, ou seja, pelos atributos intrínsecos como cor, forma, histograma, regiões homogêneas ou textura, promovendo a indexação de imagens baseada em conteúdo, como esclarece Lancaster (2004, p.213), o que justifica o número de títulos recuperado com esse enfoque na literatura da área biomédica no tocante à imagem.

Na literatura médica, diversos trabalhos abordam a criação de metadados e os diferentes padrões a serem aplicados às imagens, como também focam a discussão sobre a descrição de atributos de alto e baixo nível e sua eficácia na recuperação da informação no ambiente digital.

Em revisão de literatura sobre a informática aplicada no processamento de lâminas histopatológicas, Kothari *et al.* (2013) discorrem sobre os métodos adotados, desafios e futuras pesquisas relacionadas à tomada de decisão de diagnóstico com base na análise de imagens de lâminas. Os autores destacam os seguintes métodos: controle da qualidade da imagem capturada; extração de padrões que capturam as propriedades da imagem, como pixel, objeto e níveis semânticos; identificação de padrões (segmentação) que permitam a definição de diagnóstico ou prognóstico; e dados e visualização de informações que possibilitam o uso da imagem de lâminas histopatológicas em novas pesquisas (seleção de regiões de análise). Kothari *et al.* (2013) indicam que na etapa de análise da imagem, além da extração dos atributos de baixo nível como cor e textura, também sejam atribuídas características de alto nível com base em modelos de conhecimento (anotações).

Hsu *et al.* (2013, p. 1011) tratam do processamento semântico da informação presente na imagem médica ao destacar o gerenciamento de imagens médicas. Os autores discorrem sobre a questão da representação dos atributos de baixo nível (indexação com base no conteúdo) ser amplamente utilizada pela área, enquanto há pesquisas que realizam buscas associadas ao *Medical Subject Headings* (MeSHs), aumentando a relevância na recuperação por usar uma linguagem controlada e por serem realizadas por especialistas.

No âmbito da análise de aspectos físicos da imagem, Machado (2002) desenvolve um trabalho com o objetivo de elaborar uma modelagem conceitual no formato DICOM (*Digital Imaging and Communications in Medicine*)⁸⁹, que permite a descrição dos dados relevantes de imagens estáticas médicas para facilitar a futura recuperação. Machado (2002) propõe o armazenamento de imagens médicas e sua recuperação com base nas informações associadas à imagem (nome do paciente, data de nascimento, tipo de exame médico), nas informações obtidas a partir da interpretação (diagnóstico de patologias, órgão lesado) e nas informações agregadas por especialistas (medidas de distância e ângulos). A associação de metadados possibilita a recuperação de informações que não sejam percebidas na imagem sem a

⁸⁹ Padrão de imagem médica de diagnóstico, que estabelece série de regras e metadados para permitir comunicação entre diferentes equipamentos, desde que compatíveis com o padrão.

ingerência do especialista. Nesse sentido, a autora propõe uma solução específica no contexto da imagem médica que viabiliza a recuperação da imagem em bancos de imagens.

Carro (2003, p. 46) apresenta um modelo de metadados denominado *MedISseek* para descrever e recuperar imagens na *web*, inspirado em classes: Imagem – informações técnicas relativas à imagem digitalizada; Paciente – informações relativas ao personagem retratado; Classificação Médica – elementos que classificam as informações de acordo com propriedades da área médica; Autor – autoria da descrição; Associação – descreve relacionamento entre recursos; Marcações – conjunto de elementos que permite a anotação textual de regiões da imagem; *MedWords* – termos médicos não previstos no modelo. A estrutura em classes favorece a especialização e a implantação das informações na arquitetura RDF⁹⁰. O diferencial na proposta do autor é a incorporação da Classificação Internacional de Doenças (CID-10) na classe Classificação Médica e a previsão de inclusão de termos não existentes na CID e/ou não contemplados nos demais metadados na classe *MedWords*.

Com base no padrão *Dublin Core*⁹¹, Hasegawa e Aires (2007) também apresentam uma proposta de metadados para indexação de imagens médicas para auxiliar na recuperação da informação. Os autores destacam que os principais metadados a serem usados na indexação de imagens radiológicas são: Criador - médico ou especialista responsável pela elaboração do conteúdo; Data - data de criação da imagem; Editor - instituição responsável pelo recurso; Fonte - referência da procedência do recurso, modalidade de equipamento; Diagnóstico - resultado do diagnóstico; Patologia - nome da doença; Órgão - região do corpo em que foi realizado o exame; Posição - posição do órgão em relação a sua imagem (frontal, lateral); Tipo Imagem - indicar se a imagem é referente a Raio X, Tomografia, Ressonância Magnética, etc; Laudo - palavras-chave do laudo (HASEGAWA; AIRES, 2007, p. 52).

Analisando as propostas de metadados acima, observamos que Hasegawa e Aires (2007) apontam maior número de aspectos a serem analisados na imagem médica que Carro (2003), além de alguns aspectos com teor semelhante receberem nomeação diferenciada, como Criador e Autor. Nota-se que Carro prevê atributos relacionados diretamente à análise da imagem em questão, uma vez que destaca a classe Marcações, o que pressupõe relatar, com base na observação, a ocorrência de padrões e sua localização para posterior definição de diagnóstico. Para Hasegawa e Aires, os atributos objetivam explicitar e destacar as marcações

⁹⁰ Resource Description Framework

⁹¹ Metadados desenvolvidos com o objetivo de aperfeiçoar a organização e recuperação das informações contidas em páginas web (CUNHA; CAVALCANTI, 2008, p. 137).

ao criar aspectos descritivos específicos para Órgão e Posição. Tanto Carro (2003) como Hasegawa e Aires (2007) preveem a atribuição de palavras-chave, sendo que o primeiro autor oferece uma padronização ao adotar a lista do CID-10.

Carro (2003) e Machado (2002) incorporam o metadado Paciente na descrição da imagem. O oferecimento dessa informação requer cuidado, pois o paciente tem direito ao sigilo profissional sobre seus dados contidos em prontuários ou em exames. Nos autores estudados verifica-se a análise da imagem médica com foco na identificação e apresentação de ocorrências que se caracterizam ou não com os padrões conhecidos para definição de diagnóstico e tratamento do paciente.

Os estudos de como analisar uma imagem médica se ampliaram com a adoção de imagens digitais e seu armazenamento em bancos de dados e consequente recuperação. Na literatura, os estudos sobre bancos de imagens médicas focam o armazenamento digital de imagens; a construção de sistemas que permitem o acesso (BEZ *et al.*, 2010; PIRES *et al.*, 2004); e recuperação de imagens baseada em conteúdo a partir da extração de características das imagens, como histograma métrico (BUENO *et al.*, 2002). A organização intelectual das imagens não é evidenciada na literatura da área médica apesar do grande número de estudos sobre a análise de imagens médicas.

Em investigação sobre o desenvolvimento de um ambiente digital de organização e recuperação de imagens em saúde com foco na educação denominado Biblioteca Virtual de Imagens em Medicina (BVIM), Carrare *et al.* (2006) selecionaram 11 repositórios⁹² direcionados para a área de saúde e em plataforma *web*, e os analisaram a partir dos seguintes aspectos: imagens comercializadas; imagens com diversas subáreas; catálogo do acervo composto por metadados utilizados por bibliotecas; consulta por diretório de assunto, metadados, assunto e categoria baseada no vocabulário controlado MeSH; resultado da consulta em forma de galeria; imagens acompanhadas de informações específicas; se geram lista com as imagens selecionadas para posterior *download*; acervo composto de imagens especializadas; e se não são repositórios.

Dos repositórios investigados pelos autores (CARRARE *et al.*, 2006, p. 202), somente o *History and Special Collections* apresenta consulta por metadados, assunto e categoria. Já o *HonMedia: Images* é o único que oferece consulta por categoria baseada no vocabulário

⁹² Medical Photographic Library, Custom Medical Stock Photo, The Picture Source: science, medical and technical images, Interesting Images, HonMedia: Images, University Library: Medical Images and Illustrations, History and Special Collections, Images from the History of Medicine, MedPix: Medical Image Database, Medical Images, Images in Paediatric Cardiology.

controlado MeSH. Com base no levantamento feito pelos autores, podemos constatar que a categoria foi adotada como sendo um atributo usado na recuperação do repositório e não como um princípio a ser utilizado na organização das imagens, como um corte classificatório inicial. Mesmo assim, dos repositórios investigados, apenas dois apresentam tal funcionalidade na consulta. Esclarecemos que Carrare *et al.* (2006) não analisaram "categoria" no sentido da organização da informação, mas sim na recuperação.

Carrare *et al.* (2006, p. 203) propuseram o seguinte modelo de metadados para a BVIM: título, mídia original, espaço vetorial, dimensão, data de criação, método de obtenção, escala, descrição, padrão (colorido ou preto e branco), categoria (ilustração, radiologia ou patologia), tipo (estática ou dinâmica). Analisando o aspecto categoria proposto pelos autores (CARRARE *et al.*, 2006) observamos uma inconsistência na informação sugerida para compor o atributo. Ilustração é definida como qualquer representação de caráter não textual (CUNHA; CAVALCANTI, 2008, p. 190). Já radiologia é ramo da medicina que se dedica ao estudo e emprego dos raios X e de outras energias radiantes, e patologia é uma especialidade médica que estuda as doenças e as alterações que estas provocam no organismo. Logo, não há uma relação entre ilustração e radiologia e patologia, somente entre as duas últimas. Entendemos o emprego da expressão categoria no contexto dos autores como relativa à origem da imagem analisada.

3 USO DA IMAGEM NA PATOLOGIA

Etimologicamente, Patologia, do grego *pathos* (sofrimento, doença) e *logos* (estudo), é o estudo das doenças a partir de diferentes métodos (clínicos, bioquímicos, fisiológicos, bacteriológicos, imunológicos, entre outros). É um campo complexo que objetiva a compreensão dos mecanismos de lesões celulares e teciduais, como também os mecanismos de defesa e de reparo do organismo. Na Patologia, o material examinado pode ser analisado a partir de diferentes aspectos: causas das patologias (etiologia), mecanismos de ação (patogênese), lesões (anatomia patológica), alterações funcionais (fisiopatologia) e sinais e sintomas (propeidética) (UNIVERSIDADE FEDERAL FLUMINENSE, 2014).

O exame histopatológico consiste na análise microscópica dos tecidos para a detecção de possíveis lesões existentes, com a finalidade de informar ao clínico a natureza, a gravidade, a extensão, a evolução e a intensidade das lesões, além de sugerir ou até mesmo confirmar a causa da afecção (BRASILEIRO FILHO; BOGLIOLO, 2006).

Para Kothari *et al.* (2013, p. 1099), "a análise histopatológica é um procedimento clínico habitual para diagnosticar a presença, tipo e progressão de doenças", sendo necessária

a confecção e análise das lâminas com cortes histopatológicos para identificar padrões presentes que possibilitam a definição de diagnósticos.

Segundo Gurcan *et al.* (2009, p. 3), a análise de material histopatológico pelo patologista representa o método definitivo para a confirmação ou não da presença da doença, sua classificação e evolução. Para isso, é necessária a identificação de padrões morfológicos objetivando minimizar a subjetividade na análise e definição do diagnóstico pelo patologista. Na Patologia, o profissional deve ter conhecimento para identificar a doença como também verificar o estágio em que a mesma se encontra.

A identificação de padrões é realizada com base em amostragem quantitativa de lâminas em que ocorre determinada incidência. A localização, tamanho, forma, duração, distribuição e intensidade nas estruturas analisadas são indicadores importantes para a confirmação e a gravidade da doença. O uso de tecnologias da informação e da imagem digital na área biomédica encontra respaldo com o diagnóstico assistido por computador⁹³, atualmente foco das pesquisas em imagens biomédicas e de radiologia diagnóstica. A digitalização de lâminas histopatológicas e seu arquivamento em bancos de imagens permite a identificação de padrões morfológicos por *softwares* de reconhecimento de imagens biomédicas, agilizando a classificação da doença (GURCAN *et al.*, 2009, p. 1). O emprego da expressão "classificação" na área biomédica representa a ação de identificar uma determinada doença, a partir de padrões morfológicos.

No âmbito desse trabalho inserem-se as imagens biomédicas de lâminas histopatológicas de processos patológicos gerais confeccionadas para as aulas práticas da disciplina de Patologia Geral, do Departamento de Ciências Básicas (FCB) do Campus Universitário de Nova Friburgo da Universidade Federal Fluminense (UFF). A disciplina de Patologia Geral atende os alunos dos cursos de Biomedicina, Fonoaudiologia e Odontologia e seu ensino apresenta papel fundamental na formação de profissionais da área médica, por ser a principal conexão entre as ciências básicas e a prática clínica, assim como referência para a pesquisa básica. A produção das lâminas está inserida no projeto de ensino Utilização de Casos Clínicos e Banco de Imagens no aprendizado de Patologia Geral e aplicada no projeto de monitoria (PAES-DE-ALMEIDA, 2013), ambos desenvolvidos no âmbito da disciplina de Patologia Geral.

Além de promover a discussão de casos clínicos e de conceitos para o bom entendimento do aluno, o projeto de monitoria em questão se propõe elaborar um banco de

⁹³ Em inglês CAD (Computer-assisted diagnosis)

imagens macroscópicas e microscópicas de modo a auxiliar a autoinstrução do aluno no conteúdo teórico da disciplina de Patologia Geral. São 190 alunos de graduação diretamente beneficiados com o projeto de monitoria, sendo o conteúdo disponibilizado atualmente pelo *website* da disciplina de Patologia Geral (UNIVERSIDADE FEDERAL FLUMINENSE, 2014). As imagens oferecidas no *site* estão disponíveis para consulta no *link* Aulas, sem oferecimento da opção pesquisa por palavra-chave no acervo, não apresentando os aspectos que caracterizariam um banco de imagens, definido como “um serviço técnico de uma instituição, que seleciona, adquire, organiza, armazena e permite a recuperação de imagens fotográficas de acordo com políticas e princípios pré-estabelecidos” (RODRIGUES, 2011, p. 199).

O uso de banco de imagens como recurso didático pode ser observado na página do Departamento de Patologia Geral do Instituto de Ciências Biológicas da UFMG (UNIVERSIDADE FEDERAL DE MINAS GERAIS, 2014). O objetivo do banco de imagens é facilitar o estudo dos alunos ao promover seu acesso ao acervo de lâminas, que encontra-se organizado pelos módulos (aulas) da disciplina.

4 ASPECTOS CONCEITUAIS DA CIÊNCIA DA INFORMAÇÃO APLICADOS À ORGANIZAÇÃO DE IMAGENS BIOMÉDICAS

Organização do conhecimento, organização da informação, representação do conhecimento e representação da informação são conceitos presentes na Ciência da Informação, que interagem com o tema desse trabalho, uma vez que

A Ciência da Informação é um campo dedicado às questões científicas e à prática profissional voltadas para os problemas da efetiva comunicação do conhecimento e de seus registros entre os seres humanos, no contexto social, institucional ou individual do uso e das necessidades de informação. No tratamento dessas questões são consideradas de particular interesse as tecnologias de informação. (SARACEVIC, 1996a, p.47).

A organização e recuperação de imagens biomédicas confirmam a citação acima, uma vez que o questionamento sobre como organizar e analisar os aspectos contidos nas imagens é primordial para que a comunicação seja efetivada na prática docente no âmbito da Patologia Geral.

Conforme revelam Shera e Cleveland, a definição de Ciência da Informação estabelecida pela Conferência na *Georgia Institute of Technology*, no início dos anos 60, esclarece que

é a [disciplina] que investiga as propriedades e o comportamento da informação, as forças que regem o fluxo da informação e os meios de

processamento da informação para um máximo de acessibilidade e uso da informação. O processo inclui a origem, disseminação, coleta, organização, armazenamento, recuperação, interpretação e uso da informação. O campo deriva ou relaciona-se com a matemática, a lógica, a lingüística, a psicologia, a tecnologia, computacional, as operações de pesquisa, as artes gráficas, as comunicações, a biblioteconomia, a gestão e alguns outros campos. (SHERA; CLEVELAND, 1977, p.265).

Como observado, a área nasceu interdisciplinar, sendo a convergência de várias disciplinas como Biblioteconomia, Ciência da Informação, Ciências da Comunicação e demais áreas que fazem uso da informação na tomada de decisão, como a Biomedicina, e o acesso a dados para a definição de diagnósticos médicos. Esta interdisciplinaridade, segundo Pinheiro (2002), é um terreno vasto, onde os diferentes contextos da informação estudados pela Ciência da Informação “flutuam, correndo risco de perder seus horizontes científicos” (PINHEIRO, 2002, p. 66).

Em artigo sobre organização do conhecimento, Souza (2007) discorre sobre a abordagem histórica da Ciência da Informação, focando o trinômio documento – informação – usuário enquanto partes de um sistema de informação. Este é definido pela autora como “um conjunto de componentes apresentando uma estrutura organizada, havendo entre eles um relacionamento tal que permita ao sistema alcançar os objetivos para os quais foi criado” (SOUZA, 2007, p.112). Souza destaca que o documento e o usuário são pontos referenciais do sistema, enquanto que a informação necessita de tratamento (classificação, indexação, recuperação e disseminação) para que seja recuperada.

Para Boyce e Kraft (1985), a recuperação da informação é uma tentativa de relacionar as solicitações dos usuários com registros existentes na base de dados. Contudo, para que haja o relacionamento entre as demandas dos usuários e o resultado da busca, há de se focar o conceito de relevância. Os problemas observados na recuperação da informação mediante a adoção de máquinas mais robustas e rápidas, onde o que o ser humano concebia como relevante não era levado em conta, alimentaram os questionamentos sobre o conceito de relevância, culminando com as reflexões de Saracevic (1970). Para o autor, relevância é uma propriedade da informação, que indica uma relação entre a fonte ou emissor e o destinatário. Segundo Saracevic, “relevância é a medida de contato efetivo entre a fonte e o destinatário” (SARACEVIC, 1970, p. 112).

O conceito de relevância é voltado para a recuperação, qualidade e uso da informação, estando associado ao fornecimento de informação a tempo, de forma regular, efetiva e eficiente, capaz de eliminar a informação não relevante que, para Saracevic, é não

informação. Logo, a relevância está associada ao conceito de comunicação (PINHEIRO,1997).

Em artigo que revisa criticamente a produção científica sobre a natureza da relevância na Ciência da Informação, Saracevic (1996b) destaca quatro linhas estruturais que surgiram desde o aparecimento do conceito na literatura da área: sistemas, comunicação, contexto situacional e psicológico. O autor propõe uma quinta, denominada interativa. A concepção da natureza da relevância sob o ponto de vista de sistemas tem associação com o modelo tradicional de recuperação da informação, composto por uma série de documentos, o sistema em si e seus usuários. O conceito de relevância está relacionado com a propriedade do sistema em coletar, representar, organizar e dar como resultado documentos que atendam a busca feita pelo usuário.

No que diz respeito à comunicação, usando como modelo a teoria de informação de Shannon, Saracevic (1996b) reafirma que considera relevância como o critério a ser usado para estabelecer a efetiva comunicação entre a fonte e o destinatário. Para estabelecer essa relação, o autor considerou como elementos por parte da fonte: conhecimento do assunto, literatura sobre o assunto e arquivo do sistema, incluindo representação. Pelo lado do destinatário, considerou a estrutura cognitiva do usuário e representação, uso, contexto e valores. No contexto situacional, Saracevic aponta que os estudos dessa linha se preocupam com os usuários, com atenção para o desempenho em determinadas situações, modelos mentais de conhecimento, entre outros. O aspecto positivo dessa abordagem reside no estudo do usuário dentro de um determinado contexto informacional. Sobre a abordagem psicológica, Saracevic esclarece que consiste no estudo do estado cognitivo do pesquisador e na sua dinamicidade, o que pressupõe constantes mudanças de interpretação.

Ao propor a abordagem interativa, Saracevic (1996b) incorpora elementos das estruturas anteriores, baseada em um modelo estratificado da recuperação da informação. O autor pressupõe que a recuperação da informação requer uma interação entre computador e usuário, em diferentes níveis. O nível da superfície compreende o contato com uma interface, consistindo no ato de interagir com o sistema, fazendo buscas, por exemplo. Pelo usuário há os níveis cognitivo (interpretação, apreensão, e outros processos cognitivos), situacional (a partir de uma lacuna informacional pesquisa-se na expectativa de uma resposta) e afetivo (representa os sentimentos variados, como frustração, satisfação, entre outros). Pelo computador, destacamos conteúdo, processamento e engenharia. Saracevic propõe um sistema interdependente de relevâncias, que interagem dinamicamente entre os diferentes níveis.

O processo comunicativo entre sistema de informação e usuário exige uma reflexão e entendemos que é pertinente a inclusão de nosso problema de pesquisa, princípios teórico-metodológicos na organização e representação de imagens de lâminas histopatológicas, no contexto da Ciência da Informação, uma vez que a disciplina oferece embasamento teórico para a solução de questões surgidas nas atividades da disciplina de Patologia Geral.

5 TEORIA DA IMAGEM: CONCEITOS FUNDAMENTAIS

Diferentes estudos sobre teoria da imagem abordam como esta se constitui a partir das manifestações técnicas, estéticas e culturais. A literatura da área permite que sejam compreendidos os sistemas de representação e comunicação através das imagens.

Analisando a organização e representação de imagens a partir da abordagem semiológica, é possível vislumbrar que as imagens podem ser entendidas enquanto sistemas de significação, uma vez que a semiologia analisa o signo em relação ao seu contexto (BARTHES, 1972). Na atualidade, as diferentes tecnologias de informação e comunicação, onde os bancos de imagens estão inseridos, fazem com que os processos comunicacionais sejam objeto de estudo da semiologia.

Esclarece Barthes que “objetos, imagens, comportamentos podem significar [...] e o fazem abundantemente, mas nunca de maneira autônoma; qualquer sistema semiológico repassa-se de linguagem” (BARTHES, 1972, p.12). De acordo com o exposto pelo autor, afirma-se a necessidade da representação verbal da imagem para a transmissão da mensagem, quer seja através de uma legenda explicativa, quer seja através dos conceitos identificados em uma linguagem controlada de um sistema de informação.

A captura de uma dada realidade e a discussão teórica sobre os níveis de representação da imagem foram temas de diferentes autores. Aumont esclarece que “a produção de imagens jamais é gratuita, e, desde sempre, as imagens foram fabricadas para determinados usos, individuais ou coletivos” (AUMONT, 2008, p. 80). Sendo a imagem criada já com destinação definida, o autor destaca as três funções das imagens: o **modo simbólico**, o **modo epistêmico** e o **modo estético**. No modo simbólico, a imagem é considerada como símbolo, sendo inicialmente usada na representação religiosa por transmitir um sentimento de divindade e veneração às pessoas. O modo epistêmico destaca as informações visuais do mundo que a imagem transmite, relacionadas às coisas que nos cercam e que nos são rotineiras. Como a função de informar extrapola os aspectos não visuais da imagem, o modo epistêmico foi extremamente desenvolvido na idade moderna, com a criação de diferentes gêneros, como a imagem adotada no ensino biomédico. Com a finalidade de ser agradável ao observador e

suscitar sensações, a função estética está diretamente associada às artes ficando difícil dissociar uma da outra (AUMONT, 2008, p. 80).

Já para Sontag (2004, p. 170), a captura de uma cena pela câmera transforma aquela realidade, aquele olhar, em prova da realidade. Segundo a autora, a fotografia permite ao observador ter controle sobre o tema cuja interpretação sofrerá variações de acordo com a vivência e percepção de cada um, e se não tiver conhecimento mínimo para apreender, aquela imagem não terá significado para o espectador. A afirmação da autora explica a utilização de imagens no ensino biomédico, uma vez que o docente, ao apresentar uma lâmina com características morfológicas de uma determinada doença, objetiva levar aos discentes exemplos que os permitam identificar o diagnóstico da enfermidade que aflige o paciente.

Dubois (1993) faz uma retrospectiva das principais abordagens da teoria da fotografia e articula seu texto com base em três dimensões da fotografia: enquanto espelho do real (mimese), onde há semelhança entre a fotografia e seu referente; enquanto transformação do real (código e desconstrução), onde a fotografia é um instrumento de análise e interpretação; e como traço de um real (índice e referência), onde há a aderência do referente, nos levando a investigar a origem da imagem fotográfica.

A imagem é empregada habitualmente na comunicação por representar algo que transcende o que foi retratado. Na semiótica, essa função representativa é denominada função sígnica. Nos estudos semióticos de Peirce (2005), o autor esclarece que a divisão de signos mais importante é a referente à associação do signo em relação ao seu objeto: ícone, índice e símbolo (PEIRCE, 2005, p.64). O autor faz uma reflexão sobre a imagem no seu sentido mais geral, considerando ícones, por exemplo, expressões algébricas, cuja semelhança é auxiliada pela existência de regras convencionais. Uma vez que o objeto desta pesquisa é a análise das imagens biomédicas de lâminas histopatológicas, faremos uma leitura da obra de Peirce a partir dessa perspectiva.

Peirce insere imagem dentro da categoria ícone e fundamenta explicando que “um signo por Primeiridade é uma imagem de seu objeto e, em termos mais estritos, só pode ser uma ideia, pois deve produzir uma ideia Interpretante, e um objeto externo excita uma ideia através de uma reação sobre o cérebro” (PEIRCE, 2005, p.64). Partindo da afirmação do autor, compreende-se que a imagem de lâmina histopatológica apresenta semelhança com o que foi capturado pela câmera ou microscópio, mais precisamente, apresenta uma qualidade representativa ao observador, no nosso caso, o patologista.

Peirce acrescenta que “qualquer imagem material, como uma pintura, é grandemente convencional em seu modo de representação, porém em si mesma, sem legenda ou rótulo,

pode ser denominada hipoícone” (PEIRCE, 2005, p. 64). Os hipoícones são os signos que guardam semelhança com os objetos representados. Subdividindo o hipoícone a partir dos diferentes tipos de semelhança com o referente, Peirce obtém as classes imagem, diagrama e metáfora. A imagem guarda analogia qualitativa entre o referente e o significante. Conclui Santaella, que “uma imagem é um hipoícone porque a qualidade de sua aparência é semelhante à qualidade da aparência do objeto que a imagem representa” (SANTAELLA, 1983).

Os hipoícones, grosso modo, podem ser divididos de acordo com o modo de Primeiridade de que participem. Os que participam das qualidades simples, ou Primeira Primeiridade, são imagens; os que representam as relações, principalmente as diádicas, ou as que são assim consideradas, das partes de uma coisa através de relações em suas próprias partes, são diagramas; os que representam o caráter representativo de um representamen através da representação de um paralelismo com alguma outra coisa, são metáforas (PEIRCE, 2005, p. 64).

Apesar de na citação acima Peirce incluir as imagens como hipoícones de primeiro nível, no tocante às fotografias o autor esclarece que, mesmo que apresentem o aspecto da semelhança que as identifiquem como ícone, as imagens fotográficas estabelecem ligações por serem representações de uma realidade.

Nas fotografias, especialmente as do tipo “instantâneo”, são muito instrutivas, pois sabemos que, sob certos aspectos, são exatamente como os objetos que representam. Esta semelhança, porém, deve-se ao fato de terem sido produzidas em circunstâncias tais que foram fisicamente forçadas a corresponder ponto por ponto à natureza. Sob esse aspecto, então, as fotografias pertencem à segunda classe dos signos em virtude da conexão física (PEIRCE, 2005).

A imagem enquanto analogia de outra coisa é inserida na categoria das representações. Enquanto semelhança, a função da imagem é evocar, significar alguma coisa que não seja ela própria. No caso das lâminas histopatológicas, as imagens capturadas são representações visuais de fenômenos, sendo pertinente o estudo de como o interpretante, no nosso caso, alunos da disciplina de Patologia Geral, consegue nomear o que percebe na imagem. Inere-se que, a partir da tríade ícone-índice-símbolo, a teoria semiótica permita estudar a comunicação pela imagem na área da Patologia Geral.

6 TEORIA DA CLASSIFICAÇÃO FACETADA E DO CONCEITO COMO ABORDAGEM METODOLÓGICA PARA ORGANIZAÇÃO DE IMAGENS BIOMÉDICAS

A identificação das macrocategorias para organização do acervo de imagens de lâminas histopatológicas se apoiará metodologicamente nos princípios teóricos do modelo de conceito de Dahlberg (1978) e da Teoria da Classificação Facetada. Com base na Teoria do Conceito é possível inferir que o uso de conceitos, e não palavras, na categorização ou classificação de imagens, proporciona o controle e padronização na designação e descrição das lâminas histopatológicas, de modo a aumentar a eficácia na comunicação entre a base de dados e os alunos dos cursos de Odontologia, Biomedicina e Fonoaudiologia.

A Teoria da Classificação Facetada elaborada por Ranganathan (1967) prevê a adoção das categorias Personalidade, Matéria, Energia, Espaço e Tempo para organizar o mundo fenomenal. Adotadas na organização do conhecimento, as categorias são consideradas como representações genéricas da experiência humana, sendo formas de pensar sobre o mundo em que estamos. Segundo Langridge, “em classificação reservamos o termo ‘categoria’ para as classes mais gerais de fenômenos” (LANGRIDGE, 2006, p. 36).

No escopo deste trabalho, a adoção do termo categoria se dará em similaridade com a definição de atributo manifestada por Ranganathan, que afirma “ser qualquer propriedade ou medida qualitativa ou quantitativa de uma entidade” (RANGANATHAN, 1967, p.53, tradução nossa). O autor exemplifica sua definição, identificando os atributos de um livro: assunto, qualidade do papel, autor, idioma, ano de publicação, forma bibliográfica, encadernação, cor da capa. De maneira análoga, as lâminas histopatológicas serão analisadas de modo que seus atributos/categorias sejam identificados com fins de organização, de forma a atender à comunidade usuária. A aplicação da Teoria do Conceito e da Teoria da Classificação “reflete um modelo de abstração do mundo real, construído para determinada finalidade” (BRASCHER; CAFÉ, 2008, p. 6).

Para a definição das categorias informacionais também adotaremos a análise conceitual, que é uma etapa da indexação, quando há a identificação de conceitos que revelam o assunto do documento (*aboutness*). Especificamente na análise de imagens, Shatford (1986) propõe embasamento teórico para identificar e classificar os diferentes tipos de objetos representados em uma imagem. Compartilhamos o conceito de imagem da autora, que considera como imagens representacionais fotografias, slides, impressões ou qualquer outro item informação, bidimensional, que contenha informação imagética estática (SHATFORD, 1986, p. 39).

A análise conceitual de imagens visa possibilitar o acesso temático por parte do usuário ao acervo imagético, sem perder de vista o propósito deste acervo e seu uso. Para isso, Shatford (1986) aponta a importância em se explorar o potencial do acervo de imagens para

os possíveis futuros usuários. Essa preocupação com o potencial informativo é observada por Hjørland (1992) em artigo que afirma que a “chave” para a definição do conceito de assunto encontra-se na investigação epistemológica, no questionamento de como vamos saber o que precisamos saber sobre determinado documento para, a partir daí, descrevê-lo de forma a facilitar sua recuperação em um sistema de informação.

Shatford (1986, p. 48) propõe o desenvolvimento de uma classificação facetada para análise da imagem, adotando as facetas Quem, O que, Onde, Quando, fazendo um paralelo com as categorias ranganathianas Personalidade, Matéria, Energia, Espaço e Tempo. A primeira faceta Quem (Personalidade e/ou Matéria) corresponde aos seres ou objetos que estão na imagem. Esse questionamento se desdobra nos aspectos genérico e específico: DE quem ESPECIFICAMENTE (ou GENERICAMENTE) é essa imagem?

Correspondendo à categoria Energia, Shatford (1986, p. 52) associa a faceta O que? para determinar eventos, ações, condições e emoções dos seres ou objetos retratados na imagem. Nessa faceta observa-se além do DE GENÉRICO e DE ESPECÍFICO, o SOBRE uma vez que são traduzidas ideias abstratas ou simbólicas representadas.

Respondendo a faceta Onde (Espaço), obtêm-se áreas geográficas, espaços arquitetônicos ou locais. O aspecto DE GENÉRICO representa conceitos não específicos de um dado lugar, como por exemplo, selva, paisagem. Já o aspecto DE ESPECÍFICO identifica aspectos geográficos individuais, como cidade de Nova York, planeta Terra, etc. Já a faceta Quando (Tempo) representa tanto o tempo linear (DE ESPECÍFICO) quanto o cíclico (DE GENÉRICO). O tempo linear consiste em datas ou períodos de tempo enquanto que o tempo cíclico é representado pelas estações do ano ou parte do dia, por exemplo.

No panorama nacional, Smit (1987) destaca que a descrição de uma imagem exige orientações para nortear sua análise. Segundo a autora, “analisar uma imagem significa, quer queiramos quer não, [pressupõe] ‘traduzir’ certos elementos dessa imagem de um código icônico para um código verbal” (SMIT, 1987, p. 103). Para Smit, mesmo adotando uma linguagem documentária (artificial) para padronizar a linguagem utilizada no sistema de informação, a transcodificação da linguagem visual para a verbal não garante unanimidade na interpretação. A fim de garantir uma grande revocação, ou seja, o oferecimento de um grupo de imagens a partir do qual caberá ao usuário selecionar a imagem que melhor atenda suas demandas informacionais, Smit propõe que a descrição de imagens contemple os aspectos: “QUEM (seres vivos), ONDE (ambiente), QUANDO (tempo), ONDE (espaço), O QUE (ação) e COMO (técnica)” (SMIT, 1987, p. 109).

Smit (1996, p. 33) destaca dois questionamentos que devem ser feitos por ocasião da seleção da informação imagética: o número necessário de descritores para descrever uma imagem, de modo a atender seus usuários; e quais critérios devem-se adotar para analisar uma imagem e, conseqüentemente, identificar as categorias para sua organização.

A partir de uma amostra de imagens biomédicas de lâminas histopatológicas serão aplicadas as teorias acima – uso de conceitos e identificação de macrocategorias que representem eficazmente as imagens segundo os propósitos da disciplina de Patologia Geral e de seus discentes – para determinar os aspectos a serem apontados por ocasião da análise conceitual, de modo a definirmos a base teórica-metodológica a ser adotada na organização da informação imagética na área de Patologia Geral.

A identificação das macrocategorias e dos aspectos será validada junto aos docentes da disciplina de Patologia Geral, com o intuito de formular o embasamento teórico-metodológico a ser adotado pelos monitores e participantes do projeto de elaboração do banco de imagens de lâminas histopatológicas do Departamento de Ciências Básicas (FCB) do Campus Universitário de Nova Friburgo da Universidade Federal Fluminense (UFF).

7 CONSIDERAÇÕES FINAIS

Ao objetivar a investigação de princípios para organização de imagens de lâminas, propôs-se o desenvolvimento de princípios teórico-metodológicos para a organização desse tipo de imagem biomédica. Com a organização do acervo, o acesso da comunidade usuária da Disciplina de Patologia Geral da UFF (professores, alunos de graduação e pós-graduação) será otimizado, no que diz respeito a dar respaldo para as pesquisas macro e microscópicas dos diferentes processos patológicos, permitindo a reutilização do acervo de lâminas como recurso informacional em diferentes contextos (reuniões, encontros científicos, palestras, material didático, entre outros). O advento dos escâneres de lâminas histopatológicas proporcionou o armazenamento e disseminação em formato digital desse tipo de imagem utilizado como recurso visual nas aulas expositivas e laboratoriais de Patologia Geral.

A partir do arcabouço teórico da Ciência da Informação foi possível fazer uma reflexão sobre a organização e recuperação da informação em um dado sistema. Sendo nosso objeto de estudo a imagem de lâminas histopatológicas, a Teoria da Imagem nos forneceu embasamento conceitual para entender as informações visuais transmitidas pela imagem, e seu papel como ícone por apresentar semelhança com a realidade e com isso, possuir uma qualidade de representação do real para o patologista.

Com base nas Teorias da Classificação Facetada e do Conceito discorreu-se sobre a necessidade da identificação de categorias para a organização das imagens de lâminas

histopatológicas. Como consiste em uma pesquisa em desenvolvimento com bolsa de iniciação científica, não obtivemos até o momento uma definição sobre o critério de organização que melhor atenderá os patologistas, docentes e discentes. Observamos que a estruturação atual em módulos ou aulas oferece o critério adotado pelo professor para o desenvolvimento das aulas expositivas ou laboratoriais, não sendo possível uma recuperação dissociada do conteúdo programático em uma busca fora do contexto da disciplina. Seguindo os princípios da Teoria do Conceito e da Classificação Facetada foi possível fazer uma primeira análise do corte classificatório aplicado às imagens e das designações das lâminas no *site* da disciplina de Patologia Geral do Departamento de Ciências Básicas (FCB) do Campus Universitário de Nova Friburgo da Universidade Federal Fluminense (UFF). As imagens disponíveis foram organizadas em microscopia e macroscopia, ou seja, considerou-se como critério a ampliação do objeto a ser analisado. Dentro de cada categoria (Fotos – microscopia, Fotos – macroscopia) as imagens foram organizadas nas subclasses: Lesão reversível, Lesão irreversível, Distúrbios hemodinâmicos, Inflamação, Reparo, Alterações de crescimento, Neoplasias. As subclasses representam o conteúdo programático das aulas da disciplina de Patologia Geral, não refletindo uma análise dos fenômenos e a definição de categorias excludentes.

A adoção de um princípio para organização das imagens de lâminas histopatológicas permitirá ampliar a comunidade usuária, uma vez que estarão acessíveis a todos os pesquisadores e não somente aos alunos da disciplina de Patologia Geral.

REFERÊNCIAS

AMERICAN MEDICAL INFORMATICS ASSOCIATION. Working Group Biomedical Imaging Informatics. 2014. Disponível em: <<http://www.amia.org/programs/working-groups/biomedical-imaging-informatics>>. Acesso em: 4 de jul. 2014.

AUMONT, J. A imagem. Campinas: Papirus, 2008.

BARTHES, R. Elementos de semiologia. São Paulo: Cultrix, 1972.

BEZ, M. R. *et al.* Banco de imagens médicas para desenvolvimento de material pedagógico. In: SIMPÓSIO BRASILEIRO EM INFORMÁTICA NA EDUCAÇÃO, 21., 2010, João Pessoa. Anais... João Pessoa: UFPB, 2010. p. 1-10.

BOYCE, B.R.; KRAFT, D.H. Principles and Theories in Information Science. Annual Review of Information Science and Technology, Silver Spring, MD, USA , v. 20, p.153-178, 1985.

BRASCHER, M., CAFÉ, L. Organização da informação ou organização do conhecimento? In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 9., 2008, São Paulo. Anais... São Paulo: ANCIB, 2008.

BRASILEIRO FILHO, G.; BOGLIOLO, L. Patologia. 7. ed. Rio de Janeiro: Guanabara Koogan, 2006.

BUENO, J. M. *et al.* cbPACS: PACS com Suporte à Recuperação de Imagens Médicas Baseada em Conteúdo. In: CONGRESSO BRASILEIRO DE INFORMÁTICA EM SAÚDE, 8., 2002, Natal. Anais ... São Paulo: Sociedade Brasileira de Informática em Saúde, 2002.

CARRARE, A. P. *et al.* Uma proposta para gerenciamento e preservação de imagens em medicina na EPM/Unifesp. Ciência da Informação, Brasília, v. 35, n. 3, p. 201-208, set./dez. 2006.

CARRO, S. A. Um modelo de metadados para a indexação e recuperação de imagens médicas na web. 2003. 97 f. Dissertação (Mestrado em Computação) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2003.

CUNHA, M. B.; CAVALCANTI, C. R. O. Dicionário de Biblioteconomia e Arquivologia. Brasília: Briquet de Lemos Livros, 2008.

DAHLBERG, I. A Referent-oriented analytical concept theory of Interconcept. International Classification, Frankfurt, v. 5, n. 2, p. 142-150, 1978.

DUBOIS, P. O ato fotográfico e outros ensaios. Campinas: Papirus, 1993.

GURCAN, M.N. *et al.* Histopathological image analysis: a review. IEEE Reviews in Biomedical Engineering, [S.l.], v.2, p.147-171, 2009.

HASEGAWA, F. M.; AIRES, J. P. Proposta de um padrão de metadados para imagens médicas. Guarapuava, PR: Escola Regional de Informática, 2007.

HJORLAND, B. The concept of 'subject' in information science. Journal of Documentation, London, v. 48, n.2. p.172-200, june. 1992.

HSU, W. *et al.* Biomedical imaging informatics in the era of precision medicine: progress, challenges, and opportunities. Journal of the American Medical Informatics Association, [S.l.], v. 20, p.1010-1013, nov. 2013.

KOTHARI, S. *et al.* Pathology imaging informatics for quantitative analysis of whole-slide images. Journal of the American Medical Informatics Association, [S.l.], v. 20, p.1099-1108, nov. 2013.

LANCASTER, F. W. Indexação e resumos: teoria e prática. Brasília: Briquet de Lemos Livros, 2004.

LANGRIDGE, D. Classificação: abordagem para estudantes de Biblioteconomia. Rio de Janeiro: Interciência, 2006.

MACHADO, M. S. Modelo de metadados para armazenamento de recuperação de imagens estáticas no formato DICOM. 2002. 126 f. Dissertação (Mestrado em Computação) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2002.

PAES-DE-ALMEIDA, E. C. Utilização de casos clínicos e banco de imagens no aprendizado de Patologia Geral. Nova Friburgo, RJ, 2013. Projeto de monitoria apresentado a UFF.

PEIRCE, C. S. Semiótica. São Paulo: Perspectiva, 2005.

PINHEIRO, L.V.R.. Ciência da Informação entre sombra e luz: domínio epistemológico e campo interdisciplinar. 1997. 278 f. Tese (Doutorado em Comunicação e Cultura) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1997.

PINHEIRO, L.V.R.. Gênese da Ciência da Informação ou Sinais Anunciadores da Nova Área. In: AQUINO, M.A. O campo da Ciência da Informação: gênese, conexões e especificidades. João Pessoa: Editoria Universitária, 2002. p.61-86.

PIRES, S. R. *et al* . Banco de imagens mamográficas para treinamento na interpretação de imagens digitais. Radiologia Brasileira, São Paulo, SP, v. 37, n. 4, p. 239-244, 2004.

RANGANATHAN, S.R. Prolegomena to library classification. Bombay: Ásia Publishing House, 1967.

RODRIGUES, R. C. Análise e tematização da imagem fotográfica: determinação, delimitação e direcionamento dos discursos da imagem. 2011. 323 f. Tese (Doutorado em Ciência da Informação) – Universidade de Brasília, Brasília, 2011.

SANTAELLA, L. O que é semiótica. São Paulo: Brasiliense, 1983.

SARACEVIC, T. Ciência da Informação: origem, evolução, relações. Perspectivas em Ciência da Informação, Belo Horizonte, v.1, n. 1, p. 41-62, jan./jun. 1996a.

SARACEVIC, T. Relevance reconsidered. In: INTERNATIONAL CONFERENCE ON THE CONCEPTIONS OF LIBRARY AND INFORMATION SCIENCE, 2., 1996, Copenhagen, Denmark. Proceedings... Copenhagen, Denmark: The Royal School of Librarianship, 1996b.

SARACEVIC, T. The concept of "relevance" in Information Science: an historical review. In: SARACEVIC, T. Introduction to Information Science. New York: R. R. Bowker Co., 1970. p.111-154.

SHATFORD, S. Analyzing the subject of a picture: a theoretical approach. Cataloging and Classification Quarterly, New York, v. 6, n. 3, p. 39-62, 1986.

SHERA, J. H.; CLEVELAND, D. B. History and foundations of Information Science. Annual Review of Information Science and Technology, Washington , v. 12, p.249-275, 1977.

SMIT, J. W. A análise da imagem: um primeiro plano. In: _____. (Coord.). Análise documentária. 2. ed. Brasília: IBICT, 1987. p.99-112.

SONTAG, S. O mundo-imagem. In: _____. **Sobre fotografia**. São Paulo: Companhia das Letras, 2004. p.169-196.

SOUZA, R. F. Organização do conhecimento. In: TOUTAIN, L. M. B. B. (Org.). Para entender a ciência da informação. Salvador: EDUFBA, 2007. p.103-124.

UNIVERSIDADE FEDERAL DE MINAS GERAIS. Instituto de Ciências Biológicas. Departamento de Patologia Geral. Banco de imagem. 2014. Disponível em: <http://www.icb.ufmg.br/pat/setores/museu/banco_imagens.htm>. Acesso em: 6 mar. 2014.

UNIVERSIDADE FEDERAL FLUMINENSE. Faculdade de Odontologia de Nova Friburgo. O que é patologia, 2014. Disponível em: <<http://patogeralpunf.wix.com/punfuff#!sobre/c1t44>>. Acesso em: 6 mar. 2014.

A REPRESENTAÇÃO DA INFORMAÇÃO: CLASSIFICAÇÃO E INDEXAÇÃO AUTOMÁTICA DE DOCUMENTOS DE ARQUIVO

THE REPRESENTATION OF INFORMATION: AUTOMATIC CLASSIFICATION AND INDEXING OF ARCHIVES RECORDS

Renato Tarciso Barbosa de Sousa

Resumo: A classificação é uma função matricial do que-fazer arquivístico e um dos processos do programa de gestão de documentos. É, também, um elemento importante para a transparência e compartilhamento de informações, que são caminhos seguros para a tomada de decisão, para preservação da memória técnica e administrativa das instituições contemporâneas e para o pleno exercício da cidadania. A classificação de documentos de arquivo tem três objetivos: facilitar o acesso à informação; manter, preservar o vínculo arquivístico; fundamentar as funções arquivísticas de avaliação e de descrição. Essa importância, reconhecida pela literatura nacional e internacional, não é compatível com o nível de verticalização que se tem sobre esse tema. O pressuposto é que o arcabouço teórico-metodológico da área não consegue resolver as questões impostas pela grande massa de documentos e informações acumulada diariamente pelas organizações contemporâneas, pela disseminação das novas tecnologias e pelo desenvolvimento social. A pesquisa tem como objetivo a construção de um modelo para classificação e indexação automática de documentos de arquivo. Ela se utiliza dos métodos das Ciências Humanas e Sociais para abordar o tema a partir do material empírico privilegiado nesse trabalho, que é a literatura arquivística nacional e internacional e aquela que trata da Teoria da Classificação, da Teoria do Conceito, da Organização e Representação do Conhecimento, da Administração, da Tecnologia da Informação, da Ciência da Informação e da Diplomática Contemporânea.

Palavras-chave: Arquivística. Classificação de documentos arquivísticos. Organização de documentos de arquivo. Controle de vocabulário em arquivo. Indexação automática.

Abstract: The classification is a matrix function of the archival what-to-do one the processes of document management program. It is also un important element for transparency and information sharing, which are safety directions for making decisions, to preserve the memory of the technical and administrative contemporary institutions and the full exercise of citizenship. The classification of records has three objectives: facilitate access to information; maintain, preserve the archival bond; substantiating the archival functions of evaluation and description. This importance, recognized by national and international literature, is not compatible with the level of vertical integration we have about this topic. The assumption is that the theoretical and methodological framework of the area cannot resolve the issues imposed by the great mass of documents and information accumulated daily by contemporary organizations, the dissemination of new technologies, and social development. The research aims to build a model for automatic classification and indexing of records. It uses the methods of Humanities and Social Sciences to address the issue from the empirical field in this work, which is the national and international archival literature and the one that deals with the Classification Theory, the Theory of Concepts, Organization and Representation knowledge Management, Information Technology, Information Science and Contemporary Diplomatic.

Keywords: Archival. Classification of records. Organization of records. Vocabulary file control. Automatic indexing.

1 INTRODUÇÃO

A classificação de documentos de arquivo aparece nos estudos da área como uma preocupação há vários anos. A importância apontada pela literatura da área não teve, no mesmo nível, o estabelecimento de um instrumental teórico-metodológico consolidado. Fala-se e trabalha-se com classificação de documentos de arquivo sem agregar o desenvolvimento desse conceito proporcionado pela Filosofia, pela Teoria da Classificação e, mais recentemente, pela área de Representação e Organização da Informação. Costuma-se apresentar o caminho sem discutir os meios necessários para percorrê-lo. Importam-se, sem críticas, métodos desenvolvidos para objetos de características e natureza diferentes dos da Arquivística. E tudo isso tendo como pano de fundo a imprecisão terminológica, que parece uma marca comum da área.

A organização dos documentos tem sido um dos grandes desafios da Arquivística contemporânea. Lidar com massas tão grandes de documentos acumulados diariamente pelas instituições exige uma maior sofisticação dos esquemas de classificação. O documento arquivístico é um artefato humano com pressupostos e características específicas. O ambiente e o conteúdo são delimitados e definidos pelo sujeito acumulador, que pode ser uma pessoa física ou jurídica (organização). Ao contrário daqueles encontrados em bibliotecas, por exemplo, os documentos arquivísticos não constituem um conjunto formado em vista de uma finalidade específica: eles representam o produto da atividade do sujeito acumulador. Entender o modo como as organizações estruturam-se e como executam suas funções e atividades é compreender como os documentos são acumulados. Ele é resultado de um ato desenvolvido e, na maioria dos casos, cotidianamente repetido. A gênese se dá quando a organização tem algo a cumprir, a provar, a determinar. Surge naturalmente como resultado das ações desenvolvidas pelo sujeito acumulador. Após o registro das informações em suportes (papel, mídia magnética, microfilme, películas fotográficas, películas cinematográficas etc.), é necessário mantê-las pelos valores administrativos, técnicos, legais, fiscais, probatórios, culturais e históricos que possam conter. À medida que os documentos vão sendo acumulados, estabelecem relações entre si. Eles estão unidos por um elo criado no momento em que são produzidos e recebidos, determinado pela razão de sua elaboração e que é necessário à própria existência e a capacidade de cumprir seu objetivo. Eles são um conjunto indivisível de relações intelectuais.

A explosão do volume de documentos arquivísticos ocorreu de forma paralela com o aumento geral da informação registrada pelas sociedades humanas. Talvez, em um ritmo mais lento, mas nem por isso menos significativo. Os fenômenos da industrialização e da

urbanização, o crescimento da presença dos estados nacionais e da complexidade de seus aparatos e, secundariamente, a banalização dos meios tecnológicos de reprodução de documentos explicam em parte a afirmação anterior. A partir da década de 1980, a consolidação da microinformática fez aparecer e tornar extremamente comum nas organizações um novo tipo de documento arquivístico: o eletrônico. Não é simplesmente um novo suporte para o registro de informações originadas das funções e atividades das instituições; ele tem características próprias, que impõem a definição de modos específicos de criá-lo, classificá-lo, avaliá-lo, descrevê-lo e de atestar sua autenticidade.

O conceito de classificação de documentos de arquivo não é estranho à literatura arquivística, muito pelo contrário. Os grandes manuais sempre dedicaram capítulos para tratar do tema. Entretanto, a presença não garantiu uma verticalização teórica sobre o mesmo. A classificação, que a Lógica define como a distribuição de indivíduos em grupos distintos, de acordo com caracteres comuns e caracteres diferenciadores, aparece na literatura arquivística cindida em dois outros conceitos: classificação e arranjo.

Essas constatações dão a dimensão da necessidade de se proceder a um estudo sobre a questão, tão crucial para o profissional do arquivo e que tem sido dominada pelo empirismo e pelo senso comum. Esta parece, em alguns momentos, ser uma marca indelével da área. Silva et al. (1999, p. 208) chegam a falar de um empirismo eclético direcionado à integração das novidades tecnológicas no “velho” quadro da prática funcional de classificação, ordenação, representação descritiva e difusão de informação escrita. Há, inclusive, uma tendência de “duvidar da possibilidade de se encontrarem bases científicas para o saber arquivístico. Os seus argumentos incidem sobre o fato de estarmos perante uma prática multissecular que dispensa fundamentação teórica especial. Segundo este ponto de vista, os modelos estão criados e será na experiência, na intuição e no bom senso que se encontrarão as soluções adequadas para enfrentar a realidade.” (SILVA et al., 1999, p. 203-204).

O trabalho parte do pressuposto formulado pelo arquivista canadense Jean-Pierre Wallot, ex-presidente do Conselho Internacional de Arquivos para quem sem a pesquisa, a Arquivística está condenada a permanecer como um corpo de práticas e de receitas sem base científica. O fato de ser uma disciplina com finalidades pragmáticas a coloca em uma fronteira perigosa com o empirismo. Busca-se, ainda, com o estudo proposto, uma fundamentação teórica para o tratamento da questão da classificação de documentos de arquivo.

2 O PROBLEMA

A informação, tanto quanto os recursos humanos, materiais e financeiros, é um elemento vital para a sobrevivência de qualquer organização. Ela deve ser gerenciada de forma eficaz e eficiente, o que implica no reconhecimento de sua importância e que lhe seja dado um tratamento no mesmo nível que aquele dispensado aos outros recursos.

A organização da informação na esfera pública é uma condição para o pleno acesso à informação. O acesso deve ser visto sob dois aspectos: nível legal (*corpus* de leis) e nível técnico-científico. No nível legal, a entrada em vigor da Lei 12.527/2011 alterou essa questão no Brasil. A Lei inverteu a lógica transformando a exceção em regra, isto é, existe agora uma Lei que garante o acesso à informação pública e a exceção é o sigilo. A restrição é aquilo que pode colocar em risco a segurança do Estado e da sociedade e a privacidade das pessoas.

Entretanto, o maior entrave ao pleno acesso à informação não são as poucas restrições impostas pela lei, mas a condição de organização da informação pública. Diante disto, fica patente que os procedimentos que envolvem a produção, tratamento técnico e preservação dos acervos arquivísticos e informacionais devem estar dispostos de modo a possibilitar a sua recuperação para atender as demandas informacionais da sociedade. Aos usuários da informação devem ser franqueados todos os mecanismos que facilitem o seu acesso às informações públicas, a fim de garantir a transparência do estado.

Assim, o pleno exercício do direito à informação somente vai se realizar com a organização dos documentos públicos. Dessa forma, teremos que trazer para o cenário o conceito de gestão de documentos, que é reconhecido nacionalmente a partir da Lei 8.159/1991, a “Lei dos Arquivos no Brasil”, e internacionalmente a partir de várias normas ISO, principalmente a 15.489, a 30.300 e a 30.301.

A tecnologia da informação tem proporcionado grandes avanços nas áreas que trabalham com grandes volumes de documentos e informações. Entretanto, ela, em alguns momentos, é entendida como uma panacéia para todos os males. Em relação à classificação de documentos de arquivo, por exemplo, apesar de uma compreensão de que a tecnologia da informação, com suas várias aplicações, pode substituir completamente as rígidas estruturas de classificação por formas de acesso definidas pelos usuários.

Aos incautos devemos lembrar que o acesso, a recuperação, a busca são apenas um dos objetivos da classificação de documentos de arquivo. Autores, como Fiorella Foscarini (2010, p. 42), chegam a entender que a recuperação dos documentos é somente um benefício colateral da classificação, sendo seu propósito primário colocar os documentos individuais nos conjuntos aos quais pertencem, baseando-se no mandato e nas funções do criador, como

defende Luciana Duranti (2003, p. 43). Uma das características principais do documento de arquivo é exatamente a relação que os documentos mantêm entre si e que precisa ser preservada para dar sentido aos próprios documentos e as atividades ou ações que lhes deram origem. Inclusive, a manutenção desse contexto (representadas por essas relações entre os documentos) é um dos elementos mais destacados nas pesquisas que estudam a questão da autenticidade dos documentos digitais.

A classificação é uma função arquivística, tanto quanto a avaliação, a descrição, a criação, a preservação, a aquisição e a difusão. De acordo com Lopes (1996), o coração das práticas é formado pela classificação, avaliação e descrição. A partir desse entendimento, considero a classificação uma função matricial, pois é por meio dela que se realiza a avaliação. O instrumento arquivístico de gestão dos prazos de guarda e da destinação final (tabela de temporalidade), independente da metodologia aplicada e de sua configuração sempre tem como ponto de partida os conjuntos documentais definidos na classificação. Parte-se para a avaliação tendo como ponto inicial a classificação. No caso da descrição, a situação é semelhante. A norma internacional (ISAD-G) e a nacional (Norma brasileira de descrição arquivística, Arquivo Nacional, 2006) definem como primeiro elemento de descrição o conjunto documental, representado pelo fundo, série ou subsérie, isto é, a classificação.

Dessa forma, a classificação ganha uma envergadura, que sustenta toda a construção da prática arquivística. Ter qualidade na definição dos conjuntos documentais e de suas relações é fundamental para o que-fazer arquivístico.

A classificação é parte integrante de um Programa de Gestão de Documentos. A ISO 15.489/2001 (gestão de documentos) define sete processos de gestão de documentos, dentre eles o de organização lógica para a recuperação: classificação e indexação. A norma internacional e o E-Arq Brasil (Arquivo Nacional, 2011) indicam como instrumentos específicos e obrigatórios de gestão de documentos o plano de classificação de documentos.

Empiricamente, podemos afirmar que a excelência em situação arquivística é atestada pela existência das seguintes condições na instituição: política arquivística, existência de uma unidade político-administrativa responsável pela gestão dos documentos, de recursos humanos qualificados, de recursos materiais, de recursos financeiros e, por último, por instrumentos de gestão arquivística (plano de classificação, tabela de temporalidade etc.).

Parte-se da idéia de que o problema da classificação apresenta quatro facetas. A primeira, a indefinição e imprecisão do objeto de estudo da Arquivística. A segunda, a necessidade de um tratamento interdisciplinar da questão. A terceira, o caráter positivista das

práticas e intervenções, verificado na trajetória da produção do conhecimento na área. E, por último, a ausência de procedimentos metodológicos para a coleta de dados necessários à construção de instrumentos de classificação.

O uso cada vez mais frequente da Tecnologia da Informação possibilitou, em muitas organizações públicas e privadas, a substituição do registro e controle manual da tramitação por sistemas automatizados. É possível encontrá-los disponibilizados em redes locais e remotas. Esses sistemas têm sido desenvolvidos por profissionais de Informática. A estrutura das bases de dados, criadas para este fim, reproduz os mesmos campos definidos nas fichas, formulários e livros de protocolo.

Um dos principais problemas encontrados nos sistemas informatizados é o preenchimento do campo intitulado *Assunto* ou *Descrição do Assunto*, que tem como objetivo permitir uma identificação do conteúdo informacional do documento. Nesse campo, deve ser descrita a razão de ser do registro documental, qual o motivo de sua criação.

A análise do conteúdo desse campo nos registros de sistemas de protocolo informatizados indicou uma total falta de padrão para o preenchimento do mesmo. É impossível imaginar que uma organização, por mais complexa que ela seja, tenha uma variedade tão expressiva de assuntos. Os ganhos de eficiência no acesso à informação perdem-se num emaranhado de termos vazios, fragmentados, sem conexão e sem uma vinculação muito clara com o conteúdo informacional dos documentos. É preciso lembrar que os recursos informáticos não permitem ainda uma interpretação que resolva a imprecisão do preenchimento do campo *Assunto*. A lógica é binária.

O problema identificado é o seguinte: os dois principais campos para o acesso ao documento e à informação contida no documento de arquivo são preenchidos de forma imprecisa, tornando pouco eficiente, em termos de busca, um instrumental extremamente sofisticado, que são os sistemas utilizados nas organizações contemporâneas.

3 OBJETIVOS

3.1 Objetivo geral

O objetivo geral desta pesquisa é o de construir um modelo para a classificação de documentos de arquivo, a partir da indexação automática desses documentos mantidos em repositórios digitais, relacionando com a gestão dos prazos de guarda.

Busca-se, portanto, a construção de um modelo para a classificação de documentos de arquivo, a partir da indexação automática desses documentos. O modelo operará com um instrumento de classificação e de avaliação consolidados: o Código de Classificação de

Documentos de Arquivo das Atividades-meio e a Tabela de Temporalidade de Documentos de Arquivo das Atividades-meio, elaborados para serem aplicados nos órgãos e entidades da Administração Pública Federal. Será necessária, ainda, a construção de um vocabulário controlado para ser vinculado àquela estrutura de classificação.

Procura-se uma tentativa de solução dos problemas impostos pela organização dos arquivos. Procuramos resolver, em parte, o desafio imposto com a elaboração de um vocabulário controlado vinculado à estrutura de classificação.

Com a presença cada vez mais significativa dos documentos digitais nas instituições, avançamos nossa pesquisa para o processamento automático, vinculando diretamente classificação e avaliação de documentos. A ideia é submeter um repositório de documentos digitais a um *software* de mineração de dados. Essa mineração seria feita a partir de vocabulário controlado. A identificação dos termos desse vocabulário estaria vinculada a uma estrutura de classificação (resultado da primeira fase da pesquisa) e esta à tabela de temporalidade. Ao final do processamento teríamos um documento indexado, classificado e avaliado, isto é, com todos os elementos para seu pleno acesso e capaz de ser gerenciado arquivisticamente. Essa é a ideia.

Serão introduzidos os procedimentos micro de controle de vocabulário, que tem por objetivo o controle dos termos ou expressões. O controle de vocabulário foi inicialmente desenvolvido pela documentação e introduziu-se na Arquivística, especialmente em razão da informatização crescente dos arquivos por meio da elaboração de listas de assuntos para protocolos informatizados. Ele é um recurso para organizar e recuperar documentos e informações com consistência, gerando, conseqüentemente, confiança no sistema. É, também, um meio para produzir confiança no sistema de organização e busca de informações arquivísticas.

3.2 Objetivos específicos

- Elaborar uma metodologia para construção de vocabulário controlado para documento de arquivo;
- Relacionar o vocabulário em uma estrutura de classificação de documentos de arquivo;
- Identificar *software* de mineração de textos, que permita a mineração a partir de termos previamente selecionados;
- Analisar experiências de indexação automática;
- Elaborar modelos de tipos documentais.

4 JUSTIFICATIVA

O sucesso dessa pesquisa e dessa proposta pode resolver a importante questão da subjetividade na classificação de documentos e os problemas de qualificação de nossos recursos humanos.

Entendemos a organização dos documentos de arquivo como parte da função social do arquivista. Organizamos para possibilitar o acesso rápido, seguro e eficiente à informação. Em um primeiro momento, para a tomada segura de decisão, para a comprovação, para a garantia de direitos e deveres da instituição ou de pessoas, para um estudo retrospectivo, para a manutenção de um estoque informacional que possa servir de ponto de partida para novas atividades e, depois, para preservação do capital informacional, que permitirá um uso além das fronteiras da criação do próprio documento. Nesse sentido, a classificação assume um lugar privilegiado e aparece como ferramenta importante para esse trabalho, ou melhor, como a primeira intervenção, garantindo qualidade e os fundamentos para outras incursões no universo arquivístico, tais como: a avaliação e a descrição.

Acreditamos na possibilidade do estabelecimento de um novo marco referencial para a organização de documentos de arquivo. Esse marco, entretanto, deve ser, na nossa compreensão, o resultado do conjunto de conhecimentos cujos limites extrapolam as fronteiras da própria Arquivística. É na Filosofia, na Teoria da Classificação, na História, na Administração, na Sociologia, na Diplomática Contemporânea, na Representação e Organização do Conhecimento, na Tecnologia da Informação e na Ciência da Informação, além da Arquivística, que encontramos os fundamentos teórico-metodológicos para a operação com a organização de documentos de arquivo. Não é a adição simples de saberes, mas uma articulação a partir dos limites dos objetivos propostos.

5 METODOLOGIA

Uma pesquisa de tal envergadura requer o aporte, tanto de aspectos teórico-conceituais da Ciência da Informação e da Arquivística, quanto de métodos, técnicas e instrumentos pertencentes a estas e às disciplinas que com elas interagem, como a Teoria da Administração, a Tecnologia da Informação, dentre outras.

No âmbito da Arquivística, defende-se aqui o princípio da proveniência em toda a sua extensão, que tem como desdobramento, como diz Bellotto (2004), ou duas caras, como prefere De Luigi, ou dois graus como afirmam os canadenses, o respeito ao fundo e à ordem original. Não se perde de vista aqui que a noção de proveniência (origem dos documentos) está ligada àquela de acumulação (produção ou recebimento) e não a de recolhimento.

O respeito à ordem original, como desdobramento do princípio da proveniência, é entendido aqui não no seu sentido mais restrito. Precisamos ter outra leitura, que possibilite a manutenção da ordem original desde que ela represente, reflita a própria gênese documental e as relações orgânicas estabelecidas pelos documentos. Preservar a ordem original significa manter o documento e as suas relações orgânicas no contexto das atividades que o geraram. Pela descrição feita da realidade arquivística brasileira, percebe-se que é necessária, na maior parte dos casos, uma atuação ativa do arquivista, elaborando os instrumentos que devolvam os documentos a seu lugar de origem. Lugar intelectual e, não necessariamente, físico.

A aplicação do princípio de proveniência e de seus desdobramentos na realidade arquivística brasileira não é muito pacífica. Essa dificuldade é formada pela desorganização dos arquivos na origem (arquivos montados nos setores de trabalho e massas documentais acumuladas), pela instabilidade institucional, pela ausência de trabalhos em história administrativa e pela falta de critérios para a estruturação das organizações públicas, ou melhor, a utilização de critérios não científicos. As organizações deveriam ser estruturadas a partir de uma análise dos seus objetivos, da sua missão, e da compreensão das funções e atividades desenvolvidas. Como resultado desse estudo seria desenhado o organograma, que é a representação gráfica da estrutura organizacional e que indica a forma como o poder administrativo é dividido e delimita os níveis de autoridade, a linha de comando em seu sentido vertical. Alguns autores chegam a afirmar que, na verdade, o que existe nas organizações públicas brasileiras não é um organograma, mas sim um “personograma”. Fatores, então, de ordem política ou mesmo político-partidária e a tradição burocrática impedem que as definições ocorram daquela forma.

Entendemos como válida a utilização do princípio da proveniência e o seu desdobramento como princípios de divisão ou de classificação. Eles são compreendidos, no âmbito dessa pesquisa, como as bases que fundamentam toda a atividade de classificação. Na verdade, esses princípios são para a Teoria da Classificação a característica ou princípio da classificação ou da divisão, isto é, a qualidade ou o atributo escolhido para servir de base ao processo classificatório. Assim, existem tantas classificações quantas forem as características possíveis de serem empregadas como base da divisão.

De acordo com os fundamentos teóricos da classificação, classificar significa distribuir indivíduos em grupos distintos, de acordo com caracteres comuns e caracteres diferenciadores. Pode-se fazer essa distribuição observando-se características superficiais e mutáveis ou levando em consideração caracteres essenciais e permanentes. No primeiro caso,

a classificação é elaborada a partir de um princípio de divisão ou classificação artificial. No segundo caso, o princípio será natural.

A classificação que utiliza princípios de divisão ou de classificação artificiais não se prende à essência dos objetos distribuídos e podem ser modificados, suprimidos mesmo, sem a alteração dessa essência. É o caso, por exemplo, de distribuir os homens pela cor do cabelo, arranjar os livros em uma estante de acordo com a encadernação ou organizar os documentos arquivísticos a partir do seu suporte ou formato. Entretanto, a classificação que parte de princípios de divisão ou de classificação naturais encontra e exprime a ordem seguida pela natureza no meio das dissemelhanças quase infinitas dos seres. Ela parte da observação dos indivíduos e, em seguida, compara-os. Observando caracteres comuns em certo número de indivíduos dissemelhantes, elimina as diferenças, conserva as analogias e com estas constitui tipos de variável extensão.

Podemos, a partir dessas definições, entender a proveniência e o seu desdobramento como princípios de divisão ou de classificação naturais, pois são atributos essenciais e permanentes ao conjunto (arquivo) a ser dividido. A origem desse conjunto de documentos é uma marca indelével e inseparável do objeto e é o que lhe dá inteligibilidade, identidade e sentido. As características desse conjunto documental são delimitadas pelo sujeito acumulador (pessoa física ou jurídica) e pelas relações orgânicas que se estabelecem no interior desse conjunto.

De forma diferente ocorreu com os métodos de agrupamento dos documentos na fase anterior às determinações de Natalis de Wailly. Naquela época, os registros documentais eram reunidos a partir de características artificiais, isto é, de princípios de divisão ou de classificação baseados em características superficiais e fáceis de observar, mas que não representavam relações verdadeiras e, por esta razão, a classificação tornava-se menos perene e menos fiel às características dos objetos classificados.

A partir desse entendimento, podemos afirmar que o primeiro princípio de divisão a ser utilizado é o da proveniência, que permitirá a separação e a delimitação, dentro do patrimônio arquivístico de uma nação ou de um país, dos arquivos pertencentes a uma pessoa física ou jurídica, pública ou privada.

Enquanto o respeito aos fundos é o próprio princípio da divisão do patrimônio arquivístico de uma nação ou de um país, internamente, o princípio de respeito à ordem original indica os elementos que vão servir de base para as divisões do fundo. Esses elementos podem ser encontrados no contexto de gestação do registro documental. No sentido mais amplo, é a missão da organização que vai permitir o aparecimento dos documentos de

arquivo, mas é o desenvolvimento dessa finalidade maior que vai determinar a existência de funções, das atividades e das tarefas administrativas, que têm, nesse sentido mais restrito, como produto e subproduto o documento arquivístico. É esse processo que dá ao documento sua natureza, suas características e as suas especificidades (conforme discutido no capítulo três) e, também, seu sentido e sua própria existência. Portanto, a denominação do respeito à ordem original dada pelos canadenses, segundo grau do princípio da proveniência, parece mais adequada à compreensão.

Do mesmo modo que a identificação do fundo nem sempre é uma tarefa pacífica, os procedimentos administrativos internos que permitem o aparecimento do documento também são, em muitos casos, de difícil compreensão, pois estão, em geral, envolvidos em processos organizacionais informais, improvisados e sem registro. Defendemos aqui o estabelecimento de um ferramental teórico-metodológico para buscar a gênese documental e retirar dela as características ou atributos para fundamentar a divisão dos níveis de classificação internos do fundo de arquivo.

Este trabalho é, também, um esforço que vai além do tema proposto, pois, em todos os momentos deste estudo, parte-se da Arquivística enquanto disciplina do conhecimento humano, que, para alcançar um nível de desenvolvimento, não pode prescindir de trabalhos de pesquisa para aplicação e questionamento dos seus modelos de interpretação e de intervenção na realidade. Trabalha-se com os princípios e os conceitos de uma maneira crítica, procurando sempre tratá-los como pontos de partida para a reflexão, como o norte a ser seguido, mas de uma maneira a não limitar as possibilidades de solução aos problemas colocados. Busca-se, também, um diálogo com as áreas que possam contribuir para superar as bases intuitivas de tratamento da função classificação em Arquivística. O diálogo ocorrerá principalmente com a História, com a Filosofia, com a Teoria da Classificação, com a Administração, com a Diplomática Contemporânea, com a Tecnologia da Informação, com a Organização e Representação do Conhecimento e com a Ciência da Informação.

O percurso metodológico será apoiado em dois campos de investigação: teórico e empírico. Essa separação é apenas de fundo organizacional. A análise prescinde de uma interação e um diálogo constante entre os dois campos. Esclarecemos que o espaço empírico da pesquisa é a Administração Pública brasileira e seus instrumentos de gestão arquivística, principalmente aqueles elaborados pelo Conselho Nacional de Arquivos.

As seguintes etapas deverão ser cumpridas:

- realizar estudos para definir as bases teóricas para a elaboração de vocabulários controlados a serem aplicados em documentos de arquivo;

- analisar autores atuais da Arquivística e da Ciência da Informação que trabalham com a classificação de documentos de arquivo;
- realizar estudos a partir da Teoria da Classificação, visando estabelecer os requisitos necessários à construção de instrumentos de classificação de documentos de arquivo;
- identificar as propostas metodológicas para a elaboração de instrumentos de classificação de documentos de arquivo;
- identificar *software* de mineração de textos;
- testar os *softwares* existentes e definir quais são os requisitos necessários para o desenvolvimento da indexação a partir de uma lista selecionada de termos;
- elaborar uma metodologia para construção de vocabulário controlado em arquivo;
- definir um vocabulário controlado para as atividades-meio da Administração Pública Federal;
- relacionar o vocabulário controlado à estrutura do Código de Classificação da atividade-meio, elaborado pelo Conselho Nacional de Arquivos;
- relacionar a classificação com a avaliação de documentos de arquivo, a partir do Código de Classificação de Documentos de Arquivo e da Tabela de Temporalidade de Documentos de Arquivo, elaborados pelo Conselho Nacional de Arquivos;
- identificar um repositório digital de documentos de arquivo em um órgão da Administração Pública Federal;
- submeter os documentos digitais ao *software* minerador de textos;
- incluir os instrumentos (Código de Classificação, Tabela de Temporalidade e Vocabulário Controlado) em um sistema informatizado de gestão arquivística de documentos de arquivo.

6 CONCLUSÕES

A ideia foi a de submeter os documentos de arquivo, elaborados em meio digital, a um *software* de mineração de textos, que identificou os termos no documento a partir de um vocabulário controlado. Uma das etapas da pesquisa foi a construção desse vocabulário controlado vinculado à estrutura de classificação do Código de Classificação do Conarq. Dessa forma, quando o *software* encontrou os termos no documento, ele vinculou automaticamente a uma unidade de classificação. Com essa vinculação, será possível a determinação dos prazos de guarda desse documento. Assim, pretende-se ter, ao final do processamento, um documento de arquivo classificado, indexado e avaliado.

O protótipo foi testado em um repositório com quinhentos mil documentos (.doc, .pdf e HTML). As primeiras análises apontaram para resultados bastante satisfatórios. Os documentos foram indexados e receberam, automaticamente, um código de classificação. Entretanto, não foi possível resolver os problemas de sinonímia.

Entendemos que a grande contribuição do projeto de pesquisa em tela é a possibilidade de interrelacionar saberes de várias áreas do conhecimento, com o objetivo de um processamento técnico automatizado dos documentos de arquivo eficiente, eficaz e que gere segurança nos processos de organização, busca e avaliação desses registros documentais.

Além disso, o resultado pode criar as bases para uma metodologia para elaboração de vocabulário controlado para documentos de arquivo, o mapeamento da produção sobre classificação de documentos de arquivo e a definição de um modelo para vinculação automática entre classificação, avaliação e indexação de documentos de arquivo.

REFERÊNCIAS

ASSOCIAÇÃO DOS ARQUIVISTAS HOLANDESES. **Manual de arranjo e descrição de arquivos**. Rio de Janeiro: Arquivo Nacional, 1973.

BELLOTO, Heloísa Liberalli. **Arquivos permanentes: tratamento documental**. 2 ed. Rio de Janeiro: FGV, 2004.

BRASIL. ARQUIVO NACIONAL. **Classificação, temporalidade e destinação de documentos de arquivo relativos as atividades-meio da Administração Pública**. Rio de Janeiro: Arquivo Nacional, 2001.

BRASIL. e-ARQ Brasil: Modelo de Requisitos para Sistemas Informatizados de Gestão Arquivística de Documentos. Rio de Janeiro: Arquivo Nacional, 2011.

CAMPOS, Maria Luiza Almeida. **Em busca de princípios comuns na área de representação da informação: uma comparação entre o método de classificação facetada, o método de tesouro-baseado-em-conceito e a teoria geral da terminologia**. 1994. 207 f. Dissertação (Mestrado em Ciência da Informação). – Escola de Comunicação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1994.

CONSELHO INTERNACIONAL DE ARQUIVOS. **Norma internacional de descrição arquivística**. Rio de Janeiro: Arquivo Nacional, 2001.

COUTURE, Carol. **Les fonctions de L'Archivistique contemporaine**. Quebec: Presses de l'Université du Québec, 1999.

DÍAZ RODRIGUEZ, Alfonso. La clasificación como proceso de gestión de documentos. **Tabula**, Salamanca, n. 13, p. 79-94, 2010.

DUCHEIN, Michel. O respeito aos fundos em Arquivística: princípios teóricos e problemas práticos. **Arquivo & Administração**, Rio de Janeiro, v. 10-14, n. 1, abr. 1986.

DURANTI, Luciana, EASTWOOD, Terry, MACNEIL, Heather. **Preservation of the integrity of electronic records**. Dordrecht: Kluwer Academic Publishers, 2003.

ESTEBAN NAVARRO, Miguel Angel. La representación y la organización del conocimiento en los archivos: los lenguajes documentales ante los procesos de clasificación, ordenación y descripción. In: MARCO, Francisco Javier García (ed.). **Organización del conocimiento em sistemas de información y documentación**. Zaragoza: Librería General, 1995. p. 65-90.

FOSCARINI, Fiorella. La clasificación de documentos basada en funciones: comparación de la teoría y la práctica. **Tabula**, Salamanca, n. 13, p. 41-58, 2010.

GAGNON-ARGUIN, Louise, BANNOURI, Rabii. La classification des documents institutionnels dans les ministères du gouvernement du Québec et le défi informatique. Montreal: Conseil du Trésor, 1998.

GIL, Fernando. Classificações. In: **Enciclopédia Einaudi**. Lisboa: Imprensa Nacional Casa da Moeda, 2000. p. 90-110.

HÉON, Gilles. Les plans de classification em archivistique. **Archives**, v. 27, n. 1, p. 73-90, 1995.

LANGRIDGE, Derek. **Classificação**. Abordagem para estudantes de Biblioteconomia. Rio de Janeiro: Interciência, 1977.

LIARD, L. **Lógica**. São Paulo: Companhia Editora Nacional, 1979.

LOPES, Luis Carlos. **A informação e os arquivos**. Teorias e práticas. Niterói: Eduff, 1996.

MARTÍN-POZUELO CAMPILLOS, Maria Paz. Dificultades para la concepción de una clasificación universal em Archivística. In: MARCO, Francisco Javier García (ed.). **Organización del conocimiento en sistemas de información y documentación**. Zaragoza: Librería General, 1995. p. 91-98.

PIEIDADE, M. A. Requião. **Introdução à Teoria da Classificação**. Rio de Janeiro: Interciência, 1983.

POMBO, Olga. Da classificação dos seres à classificação dos saberes. **Leituras**. Revista da Biblioteca Nacional de Lisboa, Lisboa, n. 2, p. 19-33, 1998.

ROBERGE, Michel. La classification universelle des documents administratifs. Québec: Documentor, 1985.

SILVA, Armando Malheiro *et al.* **Arquivística**. Teoria e prática de uma ciência da informação. Porto: Afrontamento, 1999.

SOUSA, Renato Tarciso Barbosa de. **Classificação Arquivística: trajetória e apropriação de um conceito**. 2005. Tese (Doutorado em História Social) – Universidade de São Paulo.

TESSITORE, Viviane. **Arranjo**: estrutura ou função? São Paulo: Arquivo do Estado, 1989.

A NOÇÃO DE DOCUMENTAÇÃO EM MUSEUS NAS NORMAS SPECTRUM E CIDOC/ICOM

THE NOTION OF MUSEUM DOCUMENTATION AT THE SPECTRUM STANDARD AND THE CIDOC/ICOM

Juliana Monteiro
Marilda Ginez Lopes de Lara

Resumo: O objetivo deste trabalho é analisar os usos do termo *documentação* no contexto de duas normas da área de museus: as Categorias de Informação do CIDOC/ICOM e a SPECTRUM, do *Collections Trust*/Reino Unido. Os objetivos específicos são os de mapear, caracterizar e problematizar ideias, tensões e aproximações entre as duas e com as discussões teóricas da Museologia e Ciência da Informação sobre o que é documentação, sem a pretensão de se chegar a uma definição unívoca do termo. A justificativa do trabalho é a necessidade de explorar as diferenças terminológicas inerentes ao universo dos museus e discutir como operar com elas e a partir delas. A metodologia utilizada é exploratória, descritiva, comparativa e histórica. Os resultados da pesquisa reforçam a polissemia do termo documentação mostrando que a diversidade de significados influencia fortemente a prática profissional, a exemplo da ênfase, em alguns contextos, da ideia de controle de acervo dos museus com finalidade administrativa que predomina em face à finalidade de produção e difusão de conhecimento a partir das coleções. Conclui-se pela necessidade de promover a consolidação de metodologias e procedimentos que considerem a documentação simultaneamente como atividade de criação de conhecimento e prática profissional.

Palavras-chave: Documentação em museus. Museologia. Ciência da Informação. Normas.

Abstract: The present work has the objective to analyze the general uses of the term documentation within the context of the museums field: the CIDOC/ICOM Information Categories and SPECTRUM, from Collections Trust / United Kingdom. The objectives of working with these two universes are to map, characterize and discuss ideas, approaches and tensions between them and the theoretical discussions of Museology and Information Science about what is documentation, without the intention of getting any unique definition. The reason for this study is the need to explore the terminological differences inherent to the universe of museums and discuss how to work with them and from them. The methodology is exploratory, descriptive, comparative and historical. The research results reinforce the polysemy of the term documentation showing that the diversity of meanings strongly influence professional practice, such as the emphasis, in some contexts, of the idea of control of museum collection with administrative purpose that prevails in light of the aim of producing and dissemination of knowledge from collections. It concludes that it is still needed to promote the consolidation of methodologies and procedures that consider documentation as an activity of knowledge creation and also a professional practice.

Keywords: Museum Documentation. Museology. Information Science. Standards.

1 INTRODUÇÃO

O objetivo geral deste trabalho é analisar as normas SPECTRUM – Standard Procedures for Collections Recording Used in Museums / The UK Museum Collections Management Standard, lançada pela entidade britânica Collections Trust desde 1994 e as International Guidelines Object Information: the CIDOC Information Categories, do Comitê

Internacional de Documentação do Conselho Internacional de Museus – CIDOC/ICOM – em 1995. A análise será feita observando principalmente o uso do termo documentação. Para adensar a reflexão recorreremos também à literatura sobre o tema na Museologia e na Ciência da Informação.

Para identificar os traços que caracterizam o uso do termo, consideram-se as variáveis de tempo e espaço nas normas e na produção teórica da área, sem a pretensão de propor uma definição unívoca de *documentação em museus*.

Optou-se por utilizar o termo *documentação em museus* e não *documentação museológica* considerando, primeiro, que são muitas as questões ainda a discutir no próprio campo da Museologia e da Ciência da Informação, mesmo considerando o uso predominante do termo documentação museológica, no Brasil; segundo, para concentrar a pesquisa no fazer documental dos museus, que envolve o processamento dos acervos, sua gestão cotidiana e a tradução e comunicação desses conteúdos para o público.

O pressuposto teórico central para o trabalho advém da Museologia, que compreende a documentação como uma atividade fundamental para o trabalho dos museus e que envolve o processo de musealização concebida como a transformação de um *objeto* em testemunho, que possui faceta documental. O segundo pressuposto advém da Ciência da Informação, considerando a herança das discussões realizadas no âmbito da Documentação (disciplina) sobre a noção de documento. Resume-se à compreensão do documento como algo que se constrói e reconstrói a depender dos contextos informacionais, sociais, tecnológicos, institucionais e culturais envolvidos. Sendo assim, as práticas documentárias institucionalizadas, ou o fazer documental, possuem papel fundamental nesse processo de definição do potencial informativo que caracteriza o documento, logo, na definição do que é um documento em dado contexto.

O texto está dividido em duas partes: a primeira, dedicada às normas e a segunda à comparação entre as suas variações na Museologia e na Ciência da Informação. Foram selecionados autores que produzem no âmago de uma perspectiva internacionalizante, como a do CIDOC/ICOM, e de países como os do Reino Unido, Espanha e Brasil, considerando também a importância dessa instituição como promotora da primeira iniciativa internacional sobre o assunto documentação na área de museus.

Como são modelos utilizados por diferentes museus, nas conclusões procuramos pontuar e analisar comparativamente os contrastes e convergências entre prática e teoria no que se refere ao tema *documentação em museus*, a partir das principais questões conceituais levantadas nas seções anteriores. A partir disso, procuramos também propor uma visão sobre

como tais propostas podem contribuir para uma nova operacionalização da noção de documentação em museus no contexto brasileiro.

2 A SPECTRUM E AS CATEGORIAS DE INFORMAÇÃO DO CIDOC/ICOM

Para realizar a comparação entre a SPECTRUM e as Categorias de Informação do CIDOC/ICOM foram elencadas algumas categorias de análise, de modo a permitir a visualização de aproximações e distinções entre as mesmas no que tange à documentação em museus.

As duas normas datam de meados da década de 1990 – a SPECTRUM em 1994 e as Categorias do CIDOC em 1995. Ao se considerar as trajetórias históricas do CIDOC e da *Collections Trust* (então *Museum Documentation Association* ou MDA) que levaram à concepção, respectivamente, das Categorias e da SPECTRUM, pode-se inferir que os documentos nasceram em um período no qual as duas organizações expressavam o amadurecimento de discussões iniciadas entre o final da década de 1960 e início da década de 1970 em torno de padrões e formatos mínimos para a documentação em museus (MATOS, 2012).

Nos dois casos, o que se observa não é a busca de um padrão universal para os museus, o que, segundo Grant (1996), foi algo que dominou a área de museus durante muito tempo, sem frutos concretos. O que se constata nos dois documentos é a busca de padrões e formatos mínimos, com base no entendimento de que, apesar das diferenças entre as coleções, as atividades de documentação em museus necessitam de parâmetros básicos que permitam a recuperação de informação e a segurança dos objetos.

Na segunda metade da década de 1960, quando toda a área de museus se tornou cada vez mais adepta dos benefícios da informática no registro documental das coleções, o CIDOC passou a se preocupar em coordenar as experiências em andamento para criação de sistemas computadorizados para museus. Segundo Olcina (1986), por volta de 1967, o Comitê tentou criar um único sistema que pudesse abordar todos os outros, mas não teve o resultado inicialmente esperado. Já naquela época, as dificuldades de compatibilização se mostravam um desafio que chegaria até os dias atuais.

A partir desse cenário, o CIDOC iniciou um estudo para coleta e síntese de todos os sistemas espalhados ao redor do mundo, trabalho que demoraria anos para ser concluído, mas que não levou à consolidação de procedimentos que abarcassem todos os museus em sua heterogeneidade. Olcina (1986) destaca que, na mesma época no Reino Unido, o *Information Retrieval Group of the Museum Association* (IRGMA) lançou um formato mínimo com esse

objetivo – sendo este mesmo grupo a base para fundação da MDA, associação responsável pela criação da SPECTRUM. Mesmo não atingindo todas as realidades, tal formato foi adotado pelo CIDOC como base para continuação do seu trabalho de síntese de sistemas e diretrizes sobre documentação (OLCINA, 1986), o que culminaria, anos depois, nas Categorias de Informação do CIDOC/ICOM.

Nesse sentido, é preciso destacar a origem das instituições responsáveis pelo lançamento das publicações. O CIDOC/ICOM, como preconiza em seu nome, é uma instância que desde seu início se projeta como internacional – apesar da forte presença de profissionais europeus e da influência da abordagem anglo-saxã – característica esta que o próprio CIDOC tentou reverter nos anos posteriores ao lançamento das Categorias por meio da aproximação com profissionais de outros continentes, como a África (INTERNATIONAL COMMITTEE FOR DOCUMENTATION, 1996). No caso da antiga MDA/atual *Collections Trust*, a organização nasceu para consolidar discussões inicialmente afeitas ao contexto do Reino Unido, sendo que o seu objetivo de internacionalização seria estabelecido somente anos depois do lançamento da primeira versão da SPECTRUM (MATOS, 2012).

Nesse sentido, os dois documentos analisados podem ser considerados como resultados de uma construção coletiva de profissionais oriundos de diferentes museus e contextos, mas que mantêm a predominância do raciocínio europeu, particularmente anglo-saxão, sobre documentação em museus. Tal raciocínio se baseia em uma lógica destinada ao universo restrito do museu, buscando integrar as questões de controle às questões relativas à produção de informação sobre o acervo. A abordagem permanece prática em sua essência, tendo em vista o seu objetivo de funcionar como um parâmetro para o cotidiano profissional das instituições museais, ficando a documentação mais vinculada ao aspecto de organização física da coleção.

Apesar disso, em nenhum dos dois casos foi excluída a possibilidade de uso conjugado com outros padrões, tendo em vista a heterogeneidade das coleções e das instituições que as abrigam (GRANT, 1996). Em certa medida, o amadurecimento das discussões existentes desde a década de 1970, amplamente influenciadas pelo impacto da informática, trouxe para os documentos do CIDOC e da SPECTRUM o início de uma abordagem mais direcionada à relação entre gestão de acervo e gestão da informação sobre o acervo.

Em relação à atualização dos documentos, as Categorias do CIDOC não passaram por modificações desde o seu lançamento em 1995. O contrário ocorre com a SPECTRUM: na medida em que a então MDA incorporou o desejo manifesto de se aproximar de outras realidades e tornar a norma uma referência internacional, foram sendo feitas novas versões do

documento – que está agora na sua versão 4.0, lançada em 2011. As revisões realizadas ao longo dos anos incorporaram muito das sugestões e observações feitas pelos usuários da norma (MATOS, 2012), o que é possível entender como uma tentativa clara de abarcar contribuições de fora do Reino Unido.

É possível também comparar os dois modelos quanto à sua existência enquanto norma. De acordo com a *British Standards Institution* – BSI (s/d), uma norma pode ser entendida como:

Uma forma acordada, repetível de se fazer algo. É um documento que contém uma especificação técnica ou outros critérios precisos desenvolvidos para serem utilizados consistentemente como uma regra, diretriz, ou definição. As normas tornam a vida mais simples e aumentam a confiabilidade e a efetividade de muitos produtos e serviços que usamos. Elas pretendem ser uma aspiração – um resumo de boas e melhores práticas em vez de uma prática geral. As normas são criadas formando um conjunto de experiência e conhecimento de todas as partes interessadas tais como os produtores, vendedores, compradores, usuários e regulamentadores de material, produto, processo ou serviço em particular (GRUPO BSI, s/d).

A partir dessa concepção abrangente de norma, McKenna e De Loof (2009) explicam que há diferentes formas de classificar as mesmas. Uma delas é a classificação por *tipo*: as normas *de jure* e as *de fato*. As *de jure* são aquelas construídas e formalmente reconhecidas por instituições responsáveis pela regulação e estabelecimento de padrões, como a *International Standards Organization* – ISO –, ou no caso brasileiro, a Associação Brasileira de Normas Técnicas – ABNT. As normas *de fato* são aquelas não reconhecidas formalmente por nenhuma instituição, mas amplamente difundidas e aplicadas na área a qual se destina. McKenna e De Loof (2009) também classificam as normas quanto ao *contexto de sua produção*, elencando quatro categorias principais de divisão: as *in-house*, criadas para uma instituição específica; as normas de comunidade, que atendem a objetivos específicos locais; as normas nacionais, elaboradas para aplicação dentro de um país; e as normas internacionais, reconhecidas e aplicáveis a vários contextos, geralmente difundidas por instituições normativas.

Ao trazer essa discussão para a área de documentação em museus, Matos (2012) afirma que o cenário normativo da área de museus, ao contrário do que se pressupõe, não é caracterizado pela ausência de instrumentos e parâmetros, mas antes pela multiplicidade e complexidade de modelos. A multiplicidade se dá justamente pela diversidade de assuntos e de interesses de cada grupo de profissionais e instituições envolvido no desenvolvimento das normas – ora com maior foco em discutir procedimentos de gestão, estrutura de dados ou terminologia, por exemplo.

Apesar da amplitude de possibilidades de normas, Matos (2012) chama também a atenção para o fato de que essa complexidade acaba levando à fragmentação e, às vezes, ao baixo reconhecimento das normas em contextos diferentes daqueles em que originalmente foram criadas. A ausência de certificação das normas elaboradas por entidades oficiais e internacionais é quase uma constante, o que também corrobora a pouca difusão dos modelos existentes – à exceção dos esforços produzidos pelo CIDOC/ICOM e, conforme a opinião do autor, do sucesso de internacionalização da SPECTRUM.

Feitas essas considerações, é possível inserir as Categorias do CIDOC/ICOM e a SPECTRUM dentro desse cenário de diversidade normativa. No caso da SPECTRUM, o *Collections Trust* o define como norma que nasce restrita ao universo dos museus do Reino Unido, e que tem buscado nos últimos anos se internacionalizar e se consolidar como um *jeito de pensar museus* (MATOS, 2012). Contudo, a SPECTRUM também pode ser classificada como uma norma *de fato*, pois não é reconhecida oficialmente como tal por nenhuma instituição reguladora – como a ISO ou o próprio BSI. Possui ampla aceitação e reconhecimento internacional, como no caso das Categorias do CIDOC. A SPECTRUM traz orientações sobre o que fazer com o acervo – indicadas em 21 procedimentos, detalhados com fluxogramas e indicativos de grupos de informação que devem ser contemplados em cada etapa. Além disso, apresenta uma estrutura de dados que detalha os referidos grupos de informação, que subdividem em 147 unidades de informação. A diferença é que as unidades de informação da SPECTRUM não estão relacionadas somente à identificação de objetos, mas de eventos, pessoas (físicas e jurídicas) e lugares.

Em um primeiro lugar, é possível dizer que, no caso do CIDOC/ICOM, as Categorias não são definidas na publicação como um padrão, mas como um modelo de referência. Apesar disso, o documento pode ser considerado como uma norma de fato e de natureza internacional, pois como afirma Matos (2012), elas são reconhecidas como uma estrutura mínima de dados que podem ser contempladas por diferentes museus. A norma se organiza em 22 grupos de informação apenas sobre o objeto, que se subdividem em categorias de informação propriamente ditas. Em uma analogia, seria possível dizer que as categorias poderiam ser correspondentes às unidades de informação da SPECTRUM.

2.1 Os usos do termo documentação na SPECTRUM e nas Categorias do CIDOC

O glossário da SPECTRUM, que é um documento a ser consultado à parte da norma, define documentação em museus – *museum documentation* – como:

Reunião e registro de informações sobre objetos, suas histórias e sobre os processos pelos quais eles passam dentro da instituição. O propósito da

documentação é prestar contas, gerenciar e usar os objetos de modo a alcançar as metas e objetivos estabelecidos pela organização colecionadora (DAWSON, 2011, p. 1. Tradução livre).

A definição proposta pela SPECTRUM estabelece que a documentação em museus – *museum documentation* – é, ao mesmo tempo, a designação que identifica a ação de coleta e o registro de toda informação sobre os objetos, suas histórias, associações e todos os processos aos quais os objetos estão relacionados dentro do museu. Em outras palavras, a documentação é tanto o processo de organização, quanto o resultado desse processo, o que reitera a existência, de interpretação polissêmica do termo documentação.

É possível entender que, nesse cenário, a documentação foca muito mais os aspectos relativos ao controle interno da coleção do que as questões relativas ao acesso. Em certo sentido, a preponderância da faceta de controle é coerente com a definição do Glossário, apesar da sua ambiguidade: nela está posto que o primeiro propósito da documentação é o de prestar contas a respeito da coleção, ou seja, ter o controle para dar a devida transparência ao uso dos recursos empregados na preservação da mesma.

Mais um aspecto importante a ser comentado é que outro termo, gestão de coleções, aparece recorrentemente na SPECTRUM, integrando o título da norma. Apesar de não haver uma definição precisa do que é entendido por gestão de coleções na SPECTRUM, infere-se que a mesma é vista como o conjunto de procedimentos e processos pelos quais passa um objeto durante seu ciclo de vida em um museu (DAWSON; HILLHOUSE, 2011). Sendo assim, a gestão de acervos é algo que envolve não somente a identificação das peças, mas, sobretudo, a gestão da informação sobre eventos a elas relacionados, bem como pessoas, datas, lugares, etc. Também aqui é possível observar a coexistência desse entendimento ampliado sobre gestão de acervos, dentro da estrutura complexa de procedimentos e requisitos de informação sustentada pela SPECTRUM.

Embora seja necessário reconhecer o cuidado com que a SPECTRUM apresenta as conexões conceituais com sua proposta prática, fica ainda a questão a respeito do lugar da documentação. Em certa medida, interpreta-se como paradoxal, senão contraditório, o fato da documentação ser delimitada no glossário como a coleta e o registro de praticamente qualquer informação associada ao objeto dentro do museu e, na norma em si, ser tratada como algo restrito a determinado conjunto de procedimentos. A gestão de acervos acaba sendo a definição que representa a perspectiva ampliada de envolvimento de todos os processos pelos quais passa um objeto durante sua estada no museu, o que também poderia ser associado à definição de documentação ora em discussão.

A paradoxalidade da situação talvez resida na constatação de que, por mais que a SPECTRUM seja uma norma plenamente estabelecida como referência internacional de procedimentos para cuidar das coleções, não é possível refutar a inegável força de permanência desses mesmos conceitos e tradições profissionais – e sua conseqüente influência – no âmbito dessas iniciativas. Em outras palavras, compreende-se que nem todas as experiências teóricas ou práticas/profissionais que propõem o avanço em determinados aspectos têm a força de, ou desejam conscientemente, romper com parâmetros estabelecidos dentro de uma área de atuação.

Todavia, é necessário reconhecer que o objetivo principal da SPECTRUM não é promover uma análise ou crítica da noção de documentação, e sim funcionar como uma proposta para a gestão de acervo. De todo modo, é possível identificar e refletir sobre os limites da norma e seus pressupostos conceituais básicos, pois mesmo que esses não sejam diretamente amparados em alguma teoria específica, expressam, implicitamente, uma forma de conceber a documentação. Logo, é válida a análise do texto e da proposta da SPECTRUM que aqui se realiza, devido às possibilidades da convergência entre o universo da prática, que é pela norma representada, e as discussões teóricas existentes sobre documentação em museus.

Já no glossário das Categorias de Informação do CIDOC/ICOM, presentes na própria norma, a expressão utilizada é *museum documentation* – documentação de museu, em uma tradução livre. Em seu glossário, *museum documentation* significa:

Os registros que documentam a criação, o histórico, a aquisição pelo museu e a história subsequente de todos os objetos do acervo. Esses registros incluem documentos de proveniência e origem, documentos de aquisição, relatórios de conservação, registros de catalogação, imagens e pesquisas, sendo criados pela instituição detentora do objeto, proprietários anteriores, pesquisadores independentes etc. O termo também se aplica ao processo de coleta dessas informações (INTERNATIONAL COMMITTEE FOR DOCUMENTATION, 1995, p. 24. Tradução livre).

Identifica-se aqui a mesma equivalência do entendimento de documentação como processo e o resultado do processo – os registros –, tal como foi igualmente identificado no caso da SPECTRUM. A definição *museum documentation* do CIDOC/ICOM se delinea como atividade que gira em torno da coleção do museu e se infere que possua uma metodologia específica para duas ações: a coleta de informações e o gerenciamento dos documentos que lhes dão suporte – relatórios, fichas, registros diversos, imagens, etc. Não se encontra, em nenhum momento do texto, referência a uma proposta teórica ou fundamentação de base que oriente a documentação – o que, como no caso da SPECTRUM, não figura como

objetivo central do documento. Desse modo, é possível entender, considerando a essa ausência, que a documentação se encontra restrita a um conjunto de procedimentos operacionais para a identificação de um objeto.

Apesar de incluir a ideia de processo de coleta na definição, vale ressaltar que o CIDOC, nesse caso, se alonga na descrição de tipos de documentos que ‘constituem’ a documentação e não entra em detalhes sobre o que consistiria esse processo, nem como seria a metodologia adequada para a coleta de informação. A descrição dos tipos pode ser relacionada à ênfase dada aos registros / instrumentos documentários, reforçando a compreensão apresentada da documentação como resultado de um processo, ou seja, a reunião dos documentos. Essa ênfase, que inicialmente poderia passar despercebida, talvez possa ser interpretada como um indício de paradoxalidade da definição.

De modo semelhante ao caso da SPECTRUM, a característica paradoxal da definição está na verificação de que, por mais que as Categorias se configurem como um avanço na discussão sobre parâmetros para a documentação em museus dentro e fora do CIDOC, também aqui não é possível refutar a inegável força de permanência desses mesmos conceitos e tradições profissionais no âmbito do documento ora analisado. No caso da definição do CIDOC, é possível observar esse movimento de coexistência de perspectivas não necessariamente contraditórias, mas que apontam caminhos e enfoques diferentes sobre o fazer documentação em museus.

Considerando o exposto, pode-se dizer que as Categorias tratam de aspectos práticos da documentação e se propõem a especificar o tipo de informação a ser coletado no cotidiano da gestão de uma coleção, e qual o seu formato. Ressalta-se que tal consideração deve ser um ponto de atenção quando do uso das mesmas em outros contextos, nos quais a documentação nos museus possui uma abordagem diferente daquela praticada em países do universo anglo-saxão, que marca a origem das Categorias.

3 AS VARIAÇÕES EM TORNO DO TERMO DOCUMENTAÇÃO NA MUSEOLOGIA E NA CIÊNCIA DA INFORMAÇÃO

Observamos, de início, que o levantamento bibliográfico apresentado nesta seção deve ser considerado como uma contribuição inicial à discussão interdisciplinar sobre os significados atribuídos à ideia de se fazer documentação em museus. Considerando o universo pesquisado, verifica-se que o emprego do termo documentação é feito de modo diverso, pelos autores, o que reforça a constatação de que se trata de um termo polissêmico. Verifica-se, também, que o termo se encontra já profundamente enraizado na literatura da área de museus,

particularmente a produzida no âmbito da Museologia, com ocorrências também na literatura da Ciência da Informação.

3.1 O contexto europeu: Reino Unido e Espanha

É possível compreender que o uso expandido do termo documentação na área de museus se deu predominantemente a partir do contexto europeu, com protagonismo da ação de uma iniciativa internacional – o CIDOC/ICOM, criado em 1950 (STRÁNSKY, 1994). Apesar de não figurar como autora britânica ou espanhola, é necessário citar a produção da documentalista francesa Yvone Oddon, que atuou muitos anos como secretária do CIDOC/ICOM e como coordenadora do Centro de Documentação UNESCO/ICOM.

Oddon foi uma personalidade importante para a consolidação do CIDOC/ICOM como instância de discussão sobre documentação em museus, tendo sido a autora de uma obra de referência sobre o assunto (OLCINA, 1986) intitulada *Elements de documentation museographique/Elements of museum documentation*, lançada em 1968. Para Oddon, a função primeira da documentação é assegurar e preservar a identidade da coleção para facilitar a administração do museu, seguida da apresentação, interpretação e estudo do acervo. A autora entende a operacionalização da documentação do acervo em diferentes serviços – como o registro ou inventário; o catálogo; a fototeca, etc –, o que a aproxima de uma tendência organizacional característica da Biblioteconomia (CERAVOLO, 2000).

O título do livro de Oddon (1968) permite introduzir a discussão sobre o termo documentação no Reino Unido. Um fator que destaca o Reino Unido é sua histórica preocupação com as questões relacionadas à documentação dos acervos dos museus, em grande parte pela necessidade cada vez maior de prestar contas à sociedade a respeito dos recursos gastos com as coleções e com os profissionais dos museus (ROBERTS, 1985). Além disso, nos anos de 1970, foi criado o grupo *Information Retrieval Group of the Museum Association* – IRGMA – pela Associação Britânica de Museus. Esse grupo, que tinha por finalidade estudar as formas de recuperar informações sobre os acervos dos museus, logo se transformaria em uma organização chamada Associação para Documentação de Museus – em inglês, *Museum Documentation Association* / MDA – (MARÍN TORRES, 2002). Em 2008, a MDA passaria por novo processo de reestruturação, passando a se chamar *Collections Trust*.

Os autores ligados à MDA utilizam a expressão *museum documentation*, acompanhada da palavra *system*, cunhando o termo *museum documentation system* – sistema de documentação de museu. Para Light, Robert e Stewart (1986), *museum documentation system* pode ser entendido como “referente aos procedimentos usados pelos museus para gerenciar

informação relativa às suas coleções ou de relevância para suas funções curatoriais” (LIGHT; ROBERTS; STEWART, 1986, p. 1. Tradução livre). A ideia de controle da coleção é destacada pelos mesmos autores, na mesma referência, ao afirmarem que:

Os objetivos principais de tal sistema incluem auxiliar o controle e uso das coleções e assegurar a preservação de informação sobre o patrimônio cultural e ambiental. [...] Seu escopo pode englobar a documentação de coleções físicas (objetos, itens bibliográficos, material de arquivo e material audiovisual) e conjuntos de informações como detalhes de conservação e fotografias documentais, pessoas, lugares, eventos e atividades (LIGHT; ROBERTS; STEWART, 1986, p. 1. Tradução livre).

A afirmação acima mostra, primeiramente, conexão com o princípio da disciplina Documentação, conforme a concepção de Briet (2006), ao estabelecer o processamento integrado de vários tipos de documentos para uma melhor recuperação da informação; em segundo lugar, a preocupação com o gerenciamento da informação – o que será alvo de publicações específicas da MDA na década de 1990.

Na Espanha, o termo documentação de museus também aparece associado à palavra sistema, como é possível observar no título da publicação de Porta, Montserrat e Morral (1982) *Sistema de Documentación para Museos*. Nesta publicação, há uma preocupação expressa com a criação de um plano de classificação de objetos. O plano se estrutura inicialmente com a separação entre os objetos que foram feitos pelo homem e aqueles que não o foram, do que derivam agrupamentos correspondentes. Não há notas de aplicação ou explicação das classes, e o plano é apresentado como uma solução que, por inferência, deve ser adotado por todos os museus que seguirem a proposta do Departamento de Cultura da Catalunha.

O significado central atribuído à ideia de *sistema de documentación para museos* – sistema de documentação para museus – é de que a documentação se organiza como conjunto de atividades em que todas as etapas de trabalho – registro de entrada, aquisição, numeração, marcação, incorporação ao inventário geral, alocação em reserva técnica, catalogação, empréstimo e exposição – são interligadas. A interligação pressupõe uma série de procedimentos interconectados para gerenciamento das coleções. O principal objetivo do sistema é gerenciar a circulação física das peças dentro dos museus, desde o momento de sua entrada até sua localização nas áreas de guarda ou nas áreas expositivas. Em outras palavras, para auxiliar o controle da coleção – o que é o objetivo primeiro da publicação catalã.

Tal como no caso da expressão inglesa *museum documentation system*, é perceptível a preponderância da associação de aspecto prático à ideia de documentar acervos em museus, sem maior aprofundamento teórico. Uma diferença entre os dois contextos é que no caso dos

autores do Reino Unido, a concepção da ideia de sistema de documentação admite uma preocupação maior com as ações de organização da informação – relacionando eventos, datas, pessoas e lugares, às peças do acervo – e não somente com o estabelecimento de procedimentos de organização física da coleção, como é o caso da obra de Porta, Montserrat e Morral (1982). Novamente, é necessário verificar que as publicações em questão possuam objetivos distintos, mas do mesmo modo é pertinente salientar que a ocorrência dos termos permite a comparação entre seus usos e respectivos conceitos.

No final da década de 1990 e início de 2000, o que a literatura espanhola reforça é a existência de um novo posicionamento sobre a questão da documentação no âmbito dos museus. O reposicionamento, conforme esclarece Marín Torres (2002), deriva da influência do contexto anglo-saxão sobre a documentação, que a interpreta como uma atividade que faz parte de um processo maior, a gestão das coleções. Como já dito, a gestão é por esta ‘tendência’ entendida como um grupo de procedimentos destinados a aperfeiçoar as ações de organização e disponibilização das coleções institucionais, envolvendo desde técnicas de armazenagem, embalagem, transporte, conservação, avaliação, segurança e documentação. Observa-se que no âmbito dessa tendência está também a ideia de gestão da informação, que permeia transversalmente todas as técnicas e procedimentos da gestão das coleções.

Caballero Zoreda (1988) e Marín Torres (2002), por sua vez, utilizam o termo *documentación museológica* – documentação museológica – em suas publicações. Apesar disso, o entendimento não é uniforme: para Caballero Zoreda (1988), a documentação museológica é o processo de tratamento dos dados existentes nos museus, sejam eles extraídos dos objetos ou de outras fontes de informação. Caballero Zoreda (1998) cita Porta, Montserrat e Morral (1982) como uma de suas referências, com grande preocupação no detalhamento dos tipos de documento utilizados no processo. Já para Marín Torres (2002), a documentação museológica é uma parte disciplinar da ciência museológica que se encarrega do estudo da teoria, história, técnicas e procedimentos utilizados pelos museus para gerenciar e dar um sentido informativo e científico às coleções. O objeto de estudo seria a documentação museográfica propriamente dita – os registros documentais – catálogos, inventários, guias, etc.

2.1 O contexto brasileiro

No Brasil, é possível citar a existência da mesma variação em torno da palavra documentação, no contexto dos museus. Entre as variações identificadas na literatura produzida por autores da Museologia e da Ciência da Informação, estão as expressões:

catalogação (BARROSO, 1951); documentação (CAMARGO-MORO, 1986); documentação museológica (FERREZ, 1991; NASCIMENTO, 1994; LOUREIRO, 2008); documentação de museus (CERAVOLO, 1998; CERAVOLO; TÁLAMO, 2007; CERAVOLO, 2012); documentação museográfica (NOVAES, 2000) e documentação do acervo e documentação do uso do acervo (SMIT, 2010). Assim como no contexto europeu, a variação existe e indica, ora aproximações, ora afastamento entre os autores.

Entre os autores brasileiros em que é possível citar aproximações estão Ceravolo (1998), Ceravolo e Tálamo (2007) e Loureiro (2008). Mesmo usando terminologias diferentes, reconhecem que a documentação, quando realizada no âmbito dos museus, não se produz sozinha e nem possui um fim em si mesma. Para Ferrez (1991), a documentação é um sistema que permite a recuperação de qualquer informação sobre os acervos, por meio da linguagem escrita e da imagem. Autores como Smit (2010) também salientam a necessidade de se realizar as ações documentais nos museus com outra abordagem que não somente aquela afeita ao controle físico das coleções. Já Nascimento (1994) e Ceravolo (2007) sinalizam, em períodos diferentes, um aspecto importante sobre a discussão relativa ao tema no Brasil: o amadurecimento de uma base teórica e metodológica para o ato de documentar acervos de museus.

Hoje, particularmente no Brasil, a documentação é reconhecida como uma atividade que se estrutura a partir de pressupostos de estudos teóricos da Museologia. O termo predominante – documentação museológica – é um forte indicador de tal vínculo. Ou seja, enquanto prática, precisa ser orientada por conceitos como o de objeto de museu (na sua faceta documental), central para a própria área da Museologia. E se o objetivo é aproximar tal prática de aportes de outras áreas, é preciso que na sua execução sejam considerados os conceitos de documento e informação de disciplinas que a eles se dedicam, como a Documentação e a Ciência da Informação. Em outros casos, os enfoques dados pela área apenas reforçam o abandono de certas problemáticas.

Ceravolo (1998) discute a questão terminológica ao questionar a pertinência do uso da expressão documentação museológica. Para ela, não há sentido falar em documentação museológica, pois em seu entendimento a Museologia é “algo além do que a simples soma das partes das atividades de um museu” (CERAVOLO, 1998, p. 28). Segundo a autora, só faria sentido utilizar a expressão documentação museológica se

[se] empregar (ou aceitar) o domínio Museologia com foco no fato museal e não somente aos museus. Ou seja, ao se advogar a idéia de que os museus são parte de um universo muito mais amplo que ultrapassa a instituição ‘museu’ (CERAVOLO, 2012, s/p).

Em outras palavras, para a autora, o uso do termo documentação museológica, tal como se dá hoje, a insere como sinônimo de atividade exclusiva de museu. Todavia, o adjetivo ‘museológica’ qualifica a documentação como algo que pertence ou é exclusivo da Museologia – e esta disciplina se estende para além do universo restrito dos museus e de suas atividades práticas. Assim, infere-se que o uso da expressão documentação museológica contém um paradoxo, tendo em vista os termos nos quais a mesma é empregada atualmente.

Por outro lado, a mesma autora evidencia que o emprego do termo documentação museológica seria pertinente se seu significado se referisse, dentro do domínio da Museologia, às ações documentais imbricadas no processo de musealização de bens culturais. Incluir-se-iam aqui aquelas também realizadas pelos museus, mas não restrita a eles, como hoje ocorre. Ou seja, para ser ‘museológica’, a documentação deveria ser concebida dentro do sistema conceitual ou do *corpus* da Museologia, sendo-lhe própria, particular.

A crítica realizada pela autora é partilhada por este trabalho, do mesmo modo que sua afirmação sobre o fato de que o termo só tem pertinência mediante uma revisão do seu significado no âmbito do *corpus* da Museologia – o que corrobora a necessidade de aproximações e comparações com outras áreas, como a Documentação e a Ciência da Informação. Ceravolo adota a expressão documentação de museus, pois o uso do adjunto adnominal “de” traz o “sentido de alvo da ação nessa instituição e não em outra, pois particular a ela” (CERAVOLO, 1998, p. 27). Desse modo, a autora circunscreve sua reflexão apenas às ações documentais ocorridas no âmbito da musealização que os museus promovem, com a intenção de indicar suas diferentes nuances, detalhadas a seguir.

Na opinião de Ceravolo (1998), para discutir documentação de museus é preciso reconhecer duas premissas principais: que a atividade de documentar, embora parta de objetos materiais, formula operações de representação baseadas na linguagem; e que é preciso reconhecer como pressuposto básico que a concepção da documentação como um sistema orientado ao controle do acervo, não estabelece por si só uma metodologia de tratamento da informação. Portanto, conceber a documentação como um sistema não é suficiente para abarcar toda a complexidade envolvida nas operações de processamento documentário de acervos museológicos.

A autora propõe que a documentação de museus seja concebida como a reunião de dois conjuntos de atividades: o que congrega operações sobre o suporte físico das coleções – controle de localização, entrada e saída do museu, etc., ou o sistema de documentação de museus – e o conjunto de operações sobre o conteúdo das coleções – criação de padrões de indexação, classificação, etc., ou o sistema de informação documentária de museus. Ela

destaca que, no caso dos objetos resulta quase impossível, por vezes, separar o que é forma do que é conteúdo, pois ambas as dimensões estão diretamente imbricadas. Desse modo, o processo da documentação conseguiria subsidiar as atividades curatoriais, técnicas e administrativas de forma condizente com a demanda existente nas instituições.

4 CONCLUSÕES

A SPECTRUM e as Categorias de Informação do CIDOC indicam que elas perpetuam a ideia de documentação como um termo que designa um processo em que o controle da coleção é fundamental, podendo ser também o resultado do processo, ou os documentos gerados a partir dele. Desse modo, é possível dizer que as normas se alinham ao discurso predominante no contexto anglo-saxão, ou seja, refletem uma tendência pragmática. Isso permite compreender que a SPECTRUM e as Categorias do CIDOC, apesar de propostos como modelos internacionais, refletem, principalmente no que tange às definições utilizadas em seus glossários, uma perspectiva sobre o assunto que é anglo-saxônica.

Isso, entretanto, não equivale a dizer que as normas não sejam essenciais para a área de museus, tendo em vista sua abrangência e anos de aderência pelos profissionais que delas fazem uso ao redor do mundo. Seria perfeitamente possível e desejável imaginar o uso, no Brasil, das Categorias do CIDOC e da SPECTRUM – menos como soluções ou propostas acabadas sobre documentação, e mais como pontos de partida para a organização de boas práticas, realização de estudos de caso e de discussões críticas e produtivas sobre as diferenças das tendências entre os conceitos e práticas do CIDOC, do Reino Unido e do Brasil. Um desdobramento de tal raciocínio é que a consciência em torno dos limites de tais normas demanda que seu uso deva ser adaptado às realidades locais, em conjunto com outras normas produzidas dentro dos próprios países, ou pelas instituições.

Compreende-se também da análise entre as normas e da literatura consultada que a documentação em museus nasceu como um fazer prático, estruturado a partir das contribuições de profissionais atuantes nas instituições, e que se estruturou como uma atividade de perfil intermediário entre os universos teórico e prático – característica que pode ser considerada como herdada do movimento da Documentação. A documentação em museus é uma aplicação que pode extrair seus subsídios teóricos, principalmente, da disciplina Documentação e da Ciência da Informação. Essa observação se refere particularmente ao modo como tais disciplinas abordam metodologicamente as questões de representação de acervos por meio da linguagem, fornecendo subsídios para sua comunicação e interpretação. O ponto de partida para a estruturação desses sistemas ou modelos de representação é o

próprio objeto, transformado em documento no contexto dos museus, por meio das operações documentárias.

Por fim, é necessário admitir a relatividade dos significados das palavras: uma definição categórica sobre o que é documentação no contexto dos museus não faria sentido. Isso não quer dizer que não se possa realizar discussões conceituais ou práticas sobre o assunto, cotejando textos de diferentes autores, oriundos de espaços tão distintos quanto o Reino Unido, Espanha e Brasil. Para tal discussão, é necessário que se pondere sobre as especificidades linguísticas e as características da área de museus de cada lugar. Desse modo, os contornos da polissemia em torno do termo documentação no contexto dos museus, podem se tornar cada vez mais reconhecidos. Em outras palavras, os autores e profissionais que lidam com as questões documentárias em museus deverão produzir maiores avanços em questões terminológicas a partir da consciência do seu próprio lugar de fala.

Vale salientar que esta análise permite compreender que as questões aqui apresentadas não são propriamente novas, porém, é novo o foco que os pesquisadores têm dado a elas. Os problemas comentados aqui também não são novos, nova é a escala e a importância dada a eles, frente a uma realidade que exige de profissionais e teóricos de museus outra forma de lidar com questões documentárias e suas implicações e rebatimentos no cotidiano. Fica clara, sobretudo, a necessidade de se conhecer os limites dos termos, a ausência de alguns deles, as tensões inerentes ao seu uso em cada contexto e tendência identificada.

Fica clara também que a preocupação com as questões terminológicas não advém somente da necessidade dos pesquisadores e profissionais da área de museus de entenderem melhor os discursos uns dos outros, mas da própria necessidade de dar outros sentidos para a própria documentação em museus e para os acervos. Hoje, há maior consciência de que o museu tem públicos multifacetados, ávidos pela interação e dinâmicas informacionais, e com variadas opções de diversão e lazer. Esses públicos, no entanto, podem se engajar fortemente nas políticas de patrimônio cultural se sentirem que o museu faz parte de suas vidas.

REFERÊNCIAS

BRIET, S. **What is Documentation?**. English translation of the Classic French Text. DAY, R. E.; MARTINET, L.; ANGHELESCU, H. G. B. (Trad, Ed.). EUA: Scarecrow Press, 2006. 72 p.

CABALLERO ZOREDA, L. La Documentación Museológica. **Boletín de la ANABAD**, Espanha, v. 38, n. 4, p. 455-495, 1988. Disponível em: <<http://dialnet.unirioja.es/servlet/articulo?codigo=798916>>. Acesso em: 10 fev. 2013.

CAMARGO-MORO, F. de. **Museu: aquisição/documentação**. Rio de Janeiro: Livraria Eça Editora, 1986. 320 p. (Coleção Eleutherias).

CERAVOLO, S. M. Apontamentos sobre a documentação de museus. 25/12/2012. Observatório da Museologia Baiana. Disponível em: <<http://observatoriodamuseologiabaiana.blogspot.com.br>>. Acesso em: 30 abr. 2013.

CERAVOLO, S. M. **Proposta de Sistema de Informação Documentária para Museus (SIDM): a Organização da Informação para o Museu de Anatomia Veterinária da Faculdade de Medicina Veterinária e Zootecnia da Universidade da São Paulo**. São Paulo, 1998. 122f. Dissertação (Mestrado em Ciência da Informação) - Escola de Comunicações e Artes, USP, São Paulo, 1998.

CERAVOLO, S. M.; TÁLAMO, M. F. G. M. Os museus e a representação do conhecimento: uma retrospectiva sobre a documentação em museus e o processamento da informação. In: VIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO - ENANCIB , 8., 2007, Salvador. **Anais do VIII Enancib...** Salvador: PPGCI/UFBA, 2007. Disponível em: <<http://www.enancib.ppgci.ufba.br>>. Acesso em: 30 abr. 2010.

DAWNSON, A (Ed.). **Glossary for SPECTRUM 4.0**. Reino Unido: Collections Trust, 2011. Disponível em: <<http://www.collectionstrust.org.uk/spectrum/glossary-for-spectrum-4-0>>. Acesso em: 27 jun. 2014.

DAWSON, A; HILLHOUSE, S. (Ed.) **The UK Museum Collections Management**

FERREZ, H. D. Documentação museológica: teoria para uma boa prática. In: IV Fórum Nordeste de Museus, 1991, Recife, PE. **Anais...** Recife, PE: IBPC/Fundação Joaquim Nabuco, 1991, 10p. Disponível em: <<http://pt.scribd.com/doc/38689114/Documentacao-Museologica-Helena-Dodd-Ferrez>>. Acesso em: 10 mar. 2010.

GRANT, A. Museums, information and collaboration: why a single standard is not enough. In: INTERNATIONAL Council of Museums. **Study Series – International Committee for Documentation**, [S.l.], n°3,1996. Disponível em: <http://archives.icom.museum/study_series_pdf/3_ICOM-CIDOC.pdf>. Acesso em: 29 jul.2014.

GRUPO BSI. **O que é uma norma?** British Standards Institution. Brasil. Disponível em: <http://www.bsibrasil.com.br/publicacoes/sobre_normas/normas/>. Acesso em: 27 set. 2013.

INTERNATIONAL Committe for Documentation – ICOM, **CIDOC Newsletter**, [S.l.], v. 7, [s. n], 84 p., ago. 1996. Disponível em: <<http://network.icom.museum/cidoc/archives/past-newsletters/1996-newsletter/>>. Acesso em: 23 mai. 2013.

INTERNATIONAL Committee for Documentation - ICOM. **International Guidelines for Museum Object Information: the CIDOC Information Categories**. [S./l.]: ICOM, 1995. Disponível em: <<http://network.icom.museum/cidoc/resources/cidoc-standards-guidelines>>. Acesso em: 12 abr. 2014.

LIGHT, R.B.; ROBERTS, D. A.; STEWART, J.D. **Museum Documentation Systems: developments and applications**. Londres/Reino Unido: Editora Butterworths, 1986. 332 p.

LOUREIRO, J. M. M. A Documentação e suas diversas abordagens: esboço acerca da unidade museológica. In: GRANATO, M.; SANTOS, C. P. dos; LOUREIRO, M. L. N.

M.(Org.). **Documentação em Museus**. Rio de Janeiro: MAST, p. 24-32, 2008. (MAST Colloquia; 10).

MARÍN TORRES, M. T. **Historia de la documentación museológica**: la gestión de la memoria artística. Espanha: Editora Trea, 2002. 387 p.

MATOS, A. M. R. **SPECTRUM**: uma norma de gestão de coleções para os museus portugueses. Porto, 2012. 366 f. Tese (Doutoramento em Museologia) - Faculdade de Letras, Universidade do Porto, 2012. Disponível em: <<http://repositorio-aberto.up.pt/handle/10216/67304>>. Acesso em: 29 jul. 2014.

MCKENNA, G.; DE LOOF, C. **Digitisation**: standards landscape for european museums, archives, libraries. Roma/Itália: ATHENA - Working group: Identifying Standards and Developing Recommendation, 100p., 2009. Disponível em: <<http://www.athenaeurope.org/index.php?en/112/news/20/athena-booklet-digitisation-standards-landscape-for-european-museums-archives-libraries>>. Acesso em: 10 jun. 2013.

NASCIMENTO, R. A. do. Documentação Museológica e Comunicação. **Cadernos de Museologia**, Portugal, v. 3, [s.n], p. 33-43, 1994.

NOVAES, L. R. Da organização do patrimônio museológico: refletindo sobre documentação museográfica. In: SECRETARIA Municipal da Cultura de Porto Alegre. **Museologia social**. Porto Alegre: Unidade Editorial/Secretaria Municipal de Cultura, 2000. p. 43-66.

ODDON, Y. Elements of Museum Documentation: Eléments De Documentation Muséographique. In: _____. 1968. **Report on a museographical mission in Jos (Nigeria)**: February 26 - April 8,.Nigeria: Jos Museum, Centro de formação para Técnicos de Museus, 85 p. Disponível em: <<http://unesdoc.unesco.org/ulis/cgi-bin/ExtractPDF.pl?catno=158018&look=default&ll=1&display=1&lang=mb&from=&to=≥>>. Acesso em: 25 fev. 2013.

OLCINA, P. The Development and Coordination of Museum Documentation by International Agencies. In: LIGHT, Robert B.; ROBERTS, D. Andrew; STEWART, Jennifer D. **Museum Documentation Systems**: developments and applications. Londres/Reino Unido: Editora Butterworths, 1986. p. 307-314.

PORTA, E.; MONTSERRAT, R. M.; MORRAL, E. **Sistema de Documentación para Museos**. Barcelona/Espanha: INTERNATIONAL Council of Museums; Departament de Cultura de la Generalitat de Catalunya, 1982. 84 p.

SMIT, J. W. Documentação e suas diversas abordagens. In: GRANATO, M.; SANTOS, C. P. dos; LOUREIRO, M.L.N.M. (Org.). **Documentação em Museus**. Rio de Janeiro: MAST, 2008. p. 37-49. (MAST Colloquia; 10).

Standard - SPECTRUM Standard 4.0. Reino Unido: Collections Trust, 2011. Disponível em: <<http://www.collectionstrust.org.uk/spectrum>>. Acesso em: 15 jun. 2014.

STRÀNSKY, Z. Z. Object-document, or do we know what we are actually collecting?. In: **ICOFOM Study Series**, n. 23, p. 47-51, 1994. Disponível em: <<http://network.icom.museum/icofom/publications/our-publications/>>. Acesso em: 29 jul.2014.

IDENTIFICAÇÃO ARQUIVÍSTICA COMO REQUISITO DA GESTÃO DE DOCUMENTOS: CONTRIBUIÇÃO PARA APLICAÇÃO DA LEI DE ACESSO À INFORMAÇÃO NO BRASIL

Ana Célia Rodrigues

Resumo: Estudo da relação entre a identificação e a gestão de documentos, como uma contribuição para as discussões sobre os processos de aplicação da Lei de Acesso à Informação no Brasil. A identificação é um processo de trabalho arquivístico de natureza intelectual sobre o documento de arquivo e sua ligação ao órgão produtor e fundamenta as funções arquivísticas da classificação, da avaliação e da descrição, servindo também para o próprio planejamento da produção documental. Com a entrada em vigor da Lei de Acesso à Informação em maio de 2012, a demanda por documentos e informações que caracteriza o Estado Brasileiro ampliou-se cada vez mais. As conseqüências da falta de preparo da administração pública e, conseqüentemente, da dificuldade para estabelecer políticas de gestão para seus documentos e arquivos, causará impactos para a aplicabilidade da LAI no Brasil. O objetivo desta pesquisa é discutir a pertinência da identificação como requisito metodológico para o cumprimento da Lei de Acesso à Informação no âmbito de programas de gestão de documentos da administração pública estadual e municipal do Brasil. Espera-se assim, que os resultados venham contribuir com novos elementos para o debate sobre a identificação como nova função arquivística e que a sistematização de seus fundamentos teóricos e procedimentos metodológicos possam servir de referencial para os programas de gestão de documentos, cumprindo os requisitos que garantam o acesso a documentos e informações públicas no Brasil. Trata-se de projeto de pesquisa de pós-doutorado, em desenvolvimento na Universidad Carlos III de Madrid, com apoio CAPES, e na Universidade de São Paulo, que integra a produção do Grupo de Pesquisas “Gênese Documental Arquivística”, UFF/CNPq.

Palavras chave: Identificação Arquivística. Metodologia Arquivística. Diplomática. Gestão de Documentos. Lei de Acesso à Informação no Brasil.

Abstract: Study of the relationship between identification and management of documents as a contribution to the discussions on the processes of implementation of the Law on Access to Information in Brazil. Identification is a process of archival work of an intellectual nature of the document file and its connection to the national producer and founded the archival functions of classification, evaluation and description, also serving for their own planning document production. With the entry into force of the Law on Access to Information in May 2012, the demand for documents and information that characterizes the Brazilian State has expanded increasingly. The consequences of the lack of preparedness of government and, consequently, the difficulty in establishing management policies for your documents and files, cause impacts to the applicability of LAI in Brazil. The objective of this research is to discuss the relevance of identification as a methodological requirement for compliance with the Law on Access to Information under management documents of the state and municipal public administration programs in Brazil. It is expected that the results will contribute new elements to the debate on the new archival function as identification and systematization of its theoretical and methodological procedures can serve as a benchmark for document management programs, meeting the requirements guarantee access to public documents and information in Brazil. This is research project postdoctoral under development at the Universidad Carlos III de Madrid, supported CAPES, and the University of Sao Paulo, which integrates production of the Research Group "Genesis Document Archiving", UFF / CNPq.

Keywords: Identification Archivist. Methodology Archivist. Diplomatic. Document Management. Law on Access to Information in Brazil.

1 INTRODUÇÃO

Este projeto de pesquisa insere-se no campo da construção teórica de metodologias arquivísticas e pretende investigar questões ligadas à relação entre a identificação e a gestão de documentos, como uma contribuição para as discussões sobre os processos de aplicação da Lei de Acesso à Informação no Brasil.

A complexidade das ações governamentais, a crescente produção de documentos, o avanço das tecnologias de informação e a ampliação de demanda por informações que caracterizam a sociedade hoje, apontam para a necessidade de políticas públicas voltadas para a gestão de documentos, que assegure racionalidade, eficiência e transparência administrativa.

Quando se impõe a necessidade de produzir documentos autênticos, de avaliar o volume da produção documental de um órgão e de recuperar informações com a necessária agilidade exigida pela administração moderna, muitas questões se colocam e conduzem à necessidade da definição de critérios normalizados para implantar programas de gestão documental, seja em ambiente convencional ou eletrônico.

Nos anos 80 do século XX colocava-se para a arquivística o desafio de identificar imensas massas documentais acumuladas em depósitos de arquivos, os quais cresciam desordenadamente exigindo urgentes propostas de organização e avaliação para resolver o problema da superlotação dos arquivos. O termo identificação surgiu nesse contexto na Espanha designando as pesquisas desenvolvidas por grupos de arquivistas preocupados com a formulação de metodologias que solucionassem o problema da acumulação irregular de documentos. Isso contrapunha-se ao emergente desafio de implantar programas de gestão documental em países ibero-americanos. A identificação é um processo de trabalho arquivístico de natureza intelectual sobre o documento de arquivo e sua ligação ao órgão produtor.

Trata-se de uma metodologia de pesquisa que se desenvolve nos parâmetros do rigor científico e fundamenta as funções arquivísticas da classificação, da avaliação e da descrição, servindo também para o próprio planejamento da produção documental. No contexto da identificação a etapa atinente à tipologia documental encontra seus fundamentos na abordagem contemporânea da diplomática, revelando as contribuições que esta disciplina tem oferecido para a construção teórica da arquivística. O processo de produção deste conhecimento implica em reunir informações sobre o documento em seu contexto de produção e descrever os elementos capazes de revelar o “vínculo arquivístico”, que o une à

entidade produtora e aos demais documentos do mesmo conjunto ao qual ele pertence (RODRIGUES, 2008).

Na década dos 90 tem início os debates sobre a normalização no campo da arquivística impulsionados pelos processos de informatização, questões que se relacionam com a disseminação da cultura do acesso e uso dos documentos e informações neles registrada.

Neste período, assistimos a uma revisão da diplomática, cujas bases teórico-metodológicas passaram a ser utilizadas para a compreensão dos documentos contemporâneos, aplicadas aos programas de gestão de documentos e de tratamento de fundos acumulados em arquivos (DURANTI, 2005). A diplomática, revisitada pela arquivística, encontra na identificação um novo espaço para o debate científico em torno do documento de arquivo, justificando uma extensa produção científica sobre aspectos que as relacionam.

Com a consolidação da gestão de documentos no campo arquivístico, os estudos de gênese documental assumem relevância no debate teórico da área, como base para o controle normalizado do ciclo de vida dos documentos de arquivo e como condição para o estabelecimento do seu valor e acesso.

No Brasil, os dispositivos expressos na Constituição Federal de 1988, onde se afirma que "compete à administração pública, na forma da lei, a gestão da documentação governamental e as providências para franquear sua consulta a quantos dela necessitem" (§ 2º do art. 216), ofereceu os fundamentos para dotar o país de uma legislação arquivística.

A Lei nº 8.159, aprovada em 08 de janeiro de 1991, que dispõe sobre a política nacional de arquivos públicos e privados, conhecida como Lei Nacional de Arquivos, possui características notadamente conceituais, fazendo referência à gestão e ao acesso aos documentos. Evidencia os princípios federalistas e de autonomia que definem os arquivos brasileiros, estabelecendo a rede de arquivos existentes nos níveis de governo.

A gestão de documentos é conceituada na referida lei como um "conjunto de procedimentos e operações técnicas referentes à sua produção, tramitação, uso, avaliação e arquivamento em fase corrente e intermediária, visando a sua eliminação ou recolhimento para guarda permanente" (art. 3º). O art. 21 estabelece que a "legislação estadual, do Distrito Federal e municipal definirá os critérios de organização e vinculação dos arquivos, assim como a gestão e o acesso aos documentos observado o disposto na Constituição e nesta lei".

Estes princípios constitucionais fundamentaram a aprovação da Lei n. 12.527, de 18 de novembro de 2011, com o propósito de regulamentar o acesso a informações no âmbito da administração pública brasileira e seus dispositivos são aplicáveis aos três Poderes da União, Estados, Distrito Federal e Municípios.

Transparência e o acesso à informação estão previstos como direito do cidadão e dever do Estado na Constituição Federal e em diversos normativos, como a Lei de Responsabilidade Fiscal - LRF (Lei Complementar n.º 101/00), a Lei da Transparência (Lei Complementar n.º 131/09), e, mais recentemente, a Lei de Acesso à Informação - LAI (Lei n.º 12.527/11).

A Lei de Acesso à Informação vem fortalecer os instrumentos de controle da gestão pública, o combate à corrupção e a transparência dos atos administrativos. Com a entrada em vigor da LAI em maio de 2012, a demanda por documentos e informações que caracteriza o Estado Brasileiro ampliou-se cada vez mais. A lei coloca o Brasil no cenário dos 91 países, entre eles 15 da América Latina, que reconhecem o acesso amplo a documentos e informações. Todo cidadão passa agora a ter direito de solicitar cópia de qualquer informação contida em registros ou documentos produzidos ou acumulados pelos órgãos públicos.

A profusão de modelos e parâmetros conceituais usados para identificar documentos de arquivos e para planejar a gestão documental no país explica-se pela ausência de preocupações quanto à padronização, questão que não vem sendo priorizada pela área. Esta situação se agrava diante da confusão terminológica e conceitual gerada pelo uso inadequado do termo *assunto* por alguns arquivistas brasileiros e, sobretudo, da compreensão que se tem sobre ele, para reconhecer e denominar o documento de arquivo. Como resultado, observa-se uma disparidade de critérios para definir procedimentos e elaborar instrumentos que sustentem a implantação de programas de gestão documental.

A gestão de documentos é um processo arquivístico que se caracteriza como um conjunto de procedimentos aplicados no controle dos documentos durante todo o seu ciclo de vida, incidindo sobre o momento da produção (protocolo) e acumulação na primeira e segunda idade, ou seja, nos arquivos correntes e no intermediário.

Viabilizar a implantação de uma política pública de gestão de documentos pressupõe uma dupla perspectiva: dotar os órgãos públicos de instrumentos de gestão que garantam o acesso a documentos e informações e dotar o arquivo de condições técnicas e legais para seu funcionamento.

Neste cenário, algumas questões se colocam para reflexão. Como a identificação arquivística, fundamentada nos parâmetros da diplomática contemporânea, pode contribuir para a definição de requisitos metodológicos padronizados de gestão documental exigidos para o cumprimento da Lei de Acesso à Informação no Brasil? Qual a natureza dos documentos e informações conceituadas no âmbito da LAI? Como identificá-los no âmbito institucional? É possível definir padrões metodológicos de gestão documental para garantir o cumprimento da LAI?

Nesta perspectiva, o objetivo geral desta pesquisa é discutir a pertinência da identificação como requisito metodológico para o cumprimento da Lei de Acesso à Informação no âmbito de programas de gestão de documentos da administração pública estadual e municipal do Brasil. Especificamente, busca-se descrever os fundamentos teóricos e metodológicos da identificação como base para a classificação, avaliação e acesso na gestão de documentos; examinar os elementos teóricos dos conceitos de documento, informação, disponibilidade, autenticidade, integridade, gestão da informação e tratamento documental da LAI na perspectiva da diplomática e da arquivística; Sistematizar os procedimentos de identificação, classificação, avaliação e acesso e os modelos de instrumentos – formulário de identificação, plano de classificação e tabela de temporalidade – desenvolvidos para programas de gestão de documentos.

Trata-se de uma pesquisa teórica, documental e analítica que vem sendo desenvolvida a partir da revisão de literatura, sobre a identificação e a gestão de documentos, comparando os conceitos e procedimentos com os fundamentos da Lei de Acesso à Informação vigente no Brasil. Na coleta e análise de dados são utilizados também os resultados dos estudos desenvolvidos no Programa de Gestão de Documentos do Governo do Estado do Rio de Janeiro. RJ e da Prefeitura Municipal de Santos, SP, para analisar a interface da gestão de documentos e a LAI.

Trata-se de uma questão emergente no âmbito do debate científico e profissional, na busca de soluções para superar os desafios impostos para a plena aplicabilidade da Lei de Acesso à Informação.

Espera-se assim que os resultados desta pesquisa possam trazer novos elementos para o debate sobre a identificação como nova função arquivística e que a sistematização de seus fundamentos teóricos e procedimentos metodológicos possam servir de referencial para os programas de gestão de documentos, cumprindo os requisitos que garantam o acesso a documentos e informações públicas no Brasil.

2 GESTÃO DE DOCUMENTOS E ACESSO À INFORMAÇÃO: A EMERGÊNCIA DO DEBATE

A Lei de Acesso à Informação (LAI), como ficou conhecida a Lei n. 12.527/2011, define que cabem aos órgãos e entidades do poder público assegurar a “gestão transparente da informação, propiciando amplo acesso a ela e sua divulgação” e a “proteção da informação, garantindo-se sua disponibilidade, autenticidade e integridade” (art. 6, I e II).

O acesso à informação de que trata a lei, caracteriza-se como o direito de obter “orientação sobre os procedimentos para a consecução de acesso, bem como sobre o local

onde poderá ser encontrada ou obtida a informação almejada” e “informação contida em registros ou documentos, produzidos ou acumulados por seus órgãos ou entidades, recolhidos ou não a arquivos públicos.” (art. 7, I e II). Descreve os procedimentos que devem assegurar o acesso à documentos e informações, definindo o “tratamento da informação” como um “conjunto de ações referentes à produção, recepção, classificação, utilização, acesso, reprodução, transporte, transmissão, distribuição, arquivamento, armazenamento, eliminação, avaliação, destinação ou controle da informação.” (art. 3, V)

Na perspectiva da LAI, a negativa de acesso é vista como “conduta ilícita que enseja responsabilidade do agente público” se este, “utilizar indevidamente, bem como subtrair, destruir, inutilizar, desfigurar, alterar ou ocultar, total ou parcialmente, informação que se encontre sob sua guarda ou a que tenha acesso ou conhecimento em razão do exercício das atribuições de cargo, emprego ou função pública” (art. 32, II).

A LAI preconiza a transparência ativa da administração pública brasileira, principio pelo qual “as informações de interesse público deverão ser divulgadas independentemente de solicitações” (art. 3, II; art. 8) e que todos os órgãos e entidades públicas terão de divulgar anualmente uma lista com a quantidade de documentos classificados no período como reservados, secretos e ultra-secretos. (art.39)

As informações conceituadas na Lei de Acesso à Informação referem-se à:

- política, organização e serviços dos órgãos e entidades públicas.(art. 7,II e V)
- implementação, acompanhamento e resultados dos programas, projetos e ações dos órgãos e entidades públicas, bem como metas e indicadores propostos. (art. 7, VI)
- à administração do patrimônio público, utilização de recursos públicos, licitação, contratos administrativos. (art. 7, VII a)
- resultado de inspeções, auditorias, prestações e tomadas de contas realizadas pelos órgãos de controle interno e externo, incluindo prestações de contas relativas a exercícios anteriores. (art. 7, VII b)
- registros de quaisquer repasses ou transferências de recursos financeiros. (art. 8, II)
- registros das despesas. (art. 8, III)
- informações concernentes a procedimentos licitatórios, inclusive os respectivos editais e resultados, bem como a todos os contratos celebrados.(art. 8, IV)
- dados gerais para o acompanhamento de programas, ações, projetos e obras de órgãos e entidades. .(art. 8, V)
- respostas a perguntas mais frequentes da sociedade.(art. 8, VI)

A LAI fundamenta-se no direito amplo à informação, princípio constitucional de exercício da cidadania e reconhece que documentos e informações produzidas e conservadas pelo Estado brasileiro constituem bem público.

O Fórum de Direito de Acesso a Informações Públicas⁹⁴, destaca que “o documento e a informação produzida pelo agente público, pelo governante ou pelo político não pertence a ele nem ao Estado, mas sim ao cidadão” e que “o livre acesso das pessoas aos atos do governo – nos níveis municipal, estadual e federal – é um dos princípios republicanos básicos na construção de uma nação”.

É preciso lembrar que o direito à informação “vai além de um simples direito comum, trata-se de um dos direitos humanos fundamentais, que pertence a todos os indivíduos, lhes faculta exigir, receber e difundir informação”, como destaca Álvaro Baudrit Barquero (1999, p. 9)

O direito à informação ou o direito da informação para a Ciência do Direito, não é outra coisa além da faculdade que outorga o ordenamento jurídico aos indivíduos para posicionarem-se também através de gestões escritas às dependências dos órgãos públicos, a fim de solicitar informação sobre assuntos que lhe seja próprio ou de seu interesse. Direito à informação e acesso aos documentos e arquivos são conceitos muito relacionados, mas diferentes entre si; o primeiro tem a ver com a legalidade e legitimidade de obter informação; o segundo com a possibilidade material de localizar os documentos de meu interesse. (BAUDRIT BARQUERO, 1999, p. 9, *tradução nossa*)

Nesta perspectiva, direito à informação e acesso aos documentos e arquivos, são conceitos muito relacionados. Ao estar protegida a informação, “fala-se então de um direito à informação, de um direito da informação e isto não importa apenas do ponto de vista legal, mas também do ponto de vista arquivístico” (BAUDRIT BARQUERO, 1999, p. 12).

Os órgãos públicos brasileiros ainda se caracterizam por uma produção e acumulação descontrolada de documentos, situação que se agravada pela introdução das tecnologias aplicadas aos processos de trabalho. As conseqüências da falta de preparo da administração pública e, conseqüentemente, da dificuldade para estabelecer políticas de gestão para seus documentos e arquivos, causará impactos para a aplicabilidade da LAI no Brasil.

José Maria Jardim (2012), em pesquisas recentes realizadas sobre o cenário em que emerge a Lei de Acesso à Informação no Brasil, destaca que a “LAI insere o cidadão no

94 O Fórum de Direito de Acesso a Informações Públicas foi fundado em 2003 com o objetivo de agregar organizações da sociedade civil sem vínculo partidário e pressionar o governo e a sociedade pela regulamentação. Disponível em: <http://www.informacaopublica.org.br/node/2>. Acesso em 21 jul 2014.

epicentro numa ordenação jurídica à qual não corresponde uma ordenação equivalente no plano informacional” e que “sem uma agenda voltada para a definição e implementação de políticas públicas de informação, tendem a resultados limitados as iniciativas e práticas que pretendem garantir o acesso regulado pela LAI”.

Em nível federal, a Controladoria Geral da União, reconhece que estes desafios são de “natureza técnica e tecnológica e também de caráter administrativo, que incluem a necessidade de recursos financeiros e humanos - estes, devidamente capacitados - para garantir a observância do que dispõe a Lei”. (CGU, 2012). Vânia Vieira (2012), diretora de Prevenção da Corrupção da Controladoria-Geral da União, afirma que

Nós temos plena consciência de que talvez esse seja um dos principais e difíceis desafios que essa Lei de Acesso à Informação nos traz: garantir uma boa gestão da informação. Quando falo gestão, é desde o registro até o arquivamento dessa informação, porque se não tivermos um bom processo de gestão da informação, de gestão do conhecimento como um todo, por óbvio não seremos capazes de localizar, de levantar essa informação para entregá-la ao cidadão. (VIEIRA, 2012)

Em 2012, foi realizada no âmbito do Governo Federal, com apoio da Unesco, a pesquisa intitulada “Diagnóstico sobre valores, conhecimento e cultura de acesso à informação pública no Poder Executivo Federal Brasileiro⁹⁵”, com a finalidade de elaborar um diagnóstico sobre a situação enfrentada pelos servidores públicos federais no que diz respeito ao processo de produção, registro, classificação, armazenagem e disponibilização dos dados e informações sob a sua responsabilidade. Os resultados apontaram que

As práticas da administração são bastante diversas no que tange à gestão de informações. Enquanto algumas áreas já possuem sistemas de informação e bancos de dados modernos, processos digitalizados, sistemas e vários canais de comunicação com a sociedade, outras ainda estão na era do papel, enfrentando dificuldades para localizar documentos, sistematizar informações básicas e controlar o fluxo de pedidos e de respostas.(MATTA, 2012)

No âmbito federal, o Decreto que regulamenta a LAI no Executivo Federal em seu, prevê no artigo 71º, que “os órgãos e entidades adequarão suas políticas de gestão da informação, promovendo os ajustes necessários aos processos de registro, processamento, trâmite e arquivamento de documentos e informações”.

Em 16 de maio de 2012, primeiro dia de vigência da Lei de Acesso à Informação, a Controladoria-Geral da União (CGU), órgão competente para supervisionar e acompanhar os

⁹⁵ Promovida pela Controladoria-Geral da União (CGU) em parceria com a Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO), sob a coordenação do Prof. Roberto da Matta.

pedidos de acesso à informação feitos com base na LAI, registrou “até as 18 horas, 708 pedidos⁹⁶”. No dia seguinte foram registrados 1.654 pedidos⁹⁷ e após um mês, o balanço realizado pela Controladoria-Geral da União “mostra que foram feitos *mais de 10 mil pedidos* aos órgãos do poder executivo (...). Desse total, 70% teve retorno, mas nem todos correspondiam à demanda exata do cidadão. De cada 10 respondidos, apenas um era favorável à pergunta do cidadão”⁹⁸. Depois de um mês, foram mais de 10.000 solicitações, em 2013, 86.661 e até maio de 2014, foram registrados, 32.885 pedidos de informações ao governo federal⁹⁹.

No âmbito estadual, o Decreto do Governo do Estado de São Paulo que regulamenta a Lei de Acesso à Informação¹⁰⁰, destaca as interfaces entre a gestão de documentos e o acesso à informação, que se caracteriza como “aspectos inovadores em relação à Lei federal nº 12.527/2011 e às demais iniciativas (de regulamentação) divulgadas, até o momento”, ressaltam seus idealizadores. Entretanto, o Governo do Estado de São Paulo reconhece também que o “grande desafio de criar as condições e construir os mecanismos, de ordem técnica e operacional, para assegurar o seu efetivo cumprimento” (APESP, 2012).

Abrucio (2012) aponta que a LAI mostrará que é fundamental reorganizar as administrações públicas pois, “sem isso, os pedidos dos cidadãos serão frustrados, e tal cobrança será quase inútil”.

Os governos estaduais e, sobretudo, municipais terão enorme dificuldade para cumprir a nova legislação, simplesmente porque não terão muitas das informações requeridas. A origem desse problema não está na falta de organização básica dos dados de prefeituras e governadorias. Está na falta de burocracia profissional e qualificada na maior parcela da administração pública. (ABRUCIO, 2012, *grifo nosso*).

A produção desta informação que deverá ser disponibilizada ao público, ressalta Bertazzi (2011),

⁹⁶ Fonte: e-SIC/Controladoria-Geral da União. Boletim publicado em 16/05/2012. Disponível em: <http://www.cgu.gov.br/acessoainformacaogov/noticias/0102012.asp>. Acesso em: 8 ago 2012.

⁹⁷ Fonte: e-SIC/Controladoria-Geral da União. Boletim publicado em 17/05/2012. Disponível em: <http://www.cgu.gov.br/acessoainformacaogov/noticias/0102012.asp>. Acesso em: 8 ago 2012.

⁹⁸ Fonte: Jornal da Record News, publicado em 19/06/2012. Disponível em: <http://noticias.r7.com/brasil/noticias/em-apenas-um-mes-lei-de-acesso-a-a-informacao-ja-recebeu-10-mil-pedidos-20120619.html>. Acesso em: 8 ago 2012.

⁹⁹ Fonte: e-SIC/Controladoria-Geral da União. Disponível em: <http://www.cgu.gov.br/acessoainformacaogov>. Acesso em: 9 mai 2014.

¹⁰⁰ Decreto nº 58.052, de 16 de maio de 2012, que regulamenta a Lei nº 12.527, de 18 de novembro de 2011. Disponível em: http://www.arquivoestado.sp.gov.br/noticia_ver.php?id=338. Acesso em: 8 ago. 2012.

[...] demandará novos esforços e competências, novos processos e procedimentos, como por exemplo, a criação de bancos de dados e de indicadores, bem como da formatação da informação para disponibilização em diversos formatos, isso *pressupõe conhecer os documentos produzidos e as relações que mantém com outros documentos no âmbito dos procedimentos administrativos a fim de transformar em dados necessários para o atendimento da Lei de Acesso à Informação* (BERTAZZI, 2011, grifo nosso)

Estas questões remetem à necessidade de reconhecimento da natureza dos documentos e informações que circulam como produto da gestão administrativa e da necessidade de metodologias capazes de gerar tal conhecimento no âmbito da administração pública brasileira como exigência para o efetivo cumprimento da LAI, perspectiva da identificação arquivística.

Somente um programa de gestão documental bem estruturado permite o aprimoramento das rotinas administrativas e a definição dos requisitos que garantam a necessária rapidez na localização de documentos e informações exigidas para o cumprimento da legislação em vigor, conferindo eficácia e eficiência para um melhor funcionamento da administração pública em sua rotina diária de tomada de decisões, estreitando os laços entre governo e comunidades.

O desafio de implantar políticas públicas para a gestão de documentos e arquivos pode ser superado, se pautado em estratégias de cooperação. Nesta perspectiva, a universidade une-se à sociedade para debater questões novas que exigem esforços interdisciplinares entre os profissionais envolvidos com produção e tratamento de documentos e informações, em busca de referenciais teóricos e metodológicos que contribuam para o aperfeiçoamento da administração pública.

Pesquisas desenvolvidas no âmbito dos Programas de Gestão de Documentos do Governo do Estado do Rio de Janeiro¹⁰¹, RJ e da Prefeitura Municipal de Santos¹⁰², SP, vêm demonstrando a pertinência da identificação arquivística aplicada à elaboração de planos de

¹⁰¹ Projeto de pesquisa “Identificação arquivística: utilizando a diplomática contemporânea como fundamento metodológico no Programa de Gestão de Documentos do Governo do Rio de Janeiro (PGD-RJ)”, financiado pelo Programa Jovem Pesquisador UFF 2009 (PROPPi/UFF), desenvolvido através de parceria institucional entre o Arquivo Público do Estado / Casa Civil do Governo do Estado do Rio de Janeiro e o Programa de Pós-Graduação em Ciência da Informação da Universidade Federal Fluminense, sob coordenação da autora. Os instrumentos encontram-se disponíveis em: <http://www.aperj.rj.gov.br/instrumentos.htm>.

¹⁰² Ver Tabela de Temporalidade de Documentos da Secretaria Municipal de Economia e Finanças do Município de Santos, publicada no Diário Oficial do Município de Santos de 14 de fevereiro de 2012, pág. 17 a 23. Disponível em: <https://www.egov1.santos.sp.gov.br/do/0912/2012/do14022012.pdf>. Acesso em: 25 fev 2013.

classificação e tabelas de temporalidade de documentos, a partir de critérios padronizados e fundamentadas nas exigências da LAI.

No Brasil, a metodologia da identificação associada a padrões de gestão documental permanece como tema inédito no âmbito das pesquisas científicas. Além disso, a relação entre a identificação como metodologia arquivística e a Norma ISO 15489, que estabelece os requisitos dos sistemas de gestão de documentos, também não vem sendo suficientemente discutido no meio acadêmico e tampouco vem sendo aplicada pelos órgãos públicos e privados no Brasil. Essa relação, no entanto, vem encontrando relevância no debate e no modelo de gestão documental desenvolvidos na Espanha.

Estas constatações dão a dimensão da pertinência, e motivam o estudo da identificação de tipologias documentais no contexto da gestão de documentos. A sistematização dos fundamentos teóricos e metodológicos de identificação para implantar programas de gestão documental vem suprir uma lacuna que se observa no campo arquivístico brasileiro, contribuindo com a normalização de procedimentos para aplicar a Lei de Acesso à Informação no Brasil. Esses aspectos, aqui apontados, justificam o escopo e o desenvolvimento deste projeto de pesquisa.

3 A IDENTIFICAÇÃO COMO METODOLOGIA DE PESQUISA PARA A GESTÃO DE DOCUMENTOS

A identificação arquivística é uma ferramenta metodológica que se fundamenta em análise decorrente de diagnóstico elaborado sobre o documento de arquivo e seu órgão produtor, com a finalidade de propor soluções para o problema apresentado.

O termo *identificação* surge no campo da arquivística nos anos 1980, quando começou a ser utilizado na Espanha por grupos de arquivistas da Direção de Arquivos Estatais do Ministério da Cultura, para designar as tarefas de pesquisas realizadas sobre massas documentais acumuladas em arquivos, a fim de elaborar propostas de avaliação e classificação. Mas foi no âmbito do Grupo Ibero-Americano de Gestão de Documentos Administrativos, coordenado pela Espanha, e do qual participavam Espanha, Brasil, México e Portugal, que o conceito de identificação foi formulado e divulgado por Maria Luiza Conde Villaverde nas Primeiras Jornadas de Metodologia para a Identificação e Avaliação de Fundos Documentais das Administrações Públicas, realizadas em Madri, em 1991.

A partir deste momento o uso do termo e o conceito se consolidaram no meio arquivístico espanhol, sendo incorporados pelo Dicionário de Terminologia Arquivística espanhol, que a define como “fase do tratamento arquivístico que consiste na investigação e

sistematização das categorias administrativas e arquivísticas em que se sustenta a estrutura de um fundo” (DICCIONARIO, 1993, p. 37).

Neste período o Brasil integrou o movimento internacional que se formava em busca de referenciais metodológicos para resolver a superlotação dos arquivos, problema comum do qual compartilhavam os países ibero-americanos. Merecem referencia os trabalhos dos grupos que se criaram no âmbito do Programa de Modernização Administrativa do Arquivo Nacional¹⁰³, cujos resultados permitiram a formulação do conceito de identificação incorporado pelo Dicionário Brasileiro de Terminologia Arquivística (2005) que considera a *identificação* como o “processo de reconhecimento, sistematização e registro de informações sobre arquivos, com vistas ao seu controle físico e/ou intelectual”.

Estudos desenvolvidos anteriormente em Rodrigues (2003) sobre a construção teórica da metodologia de identificação aplicada à arquivística permitiram sistematizar os aspectos teóricos e metodológicos que a caracterizam e sua pertinência como ferramenta de pesquisa para o desenvolvimento das práticas arquivísticas, relacionando-a com as funções de classificação, avaliação, descrição e também com o planejamento da produção documental. (Rodrigues, 2008, 2012; 2013).

Trata-se de uma tarefa de pesquisa sobre os elementos que caracterizam os dois objetos de estudos da identificação: órgão produtor, analisando o elemento orgânico (estrutura administrativa) e o elemento funcional (competências, funções e atividades) e tipologia documental, estudo que se realiza com base no reconhecimento dos elementos externos, que se referem a estrutura física, a forma de apresentação do documento (gênero, suporte, formato e forma) e internos, o “conteúdo substantivo do documento (atividade) e natureza de sua proveniência e função” (BELLOTTO, 2004), para denominar o tipo e definir a série documental.

O procedimento de identificação parte do conceito de documento, dos elementos constitutivos da sua estrutura, a relação do documento com o ato, dos elementos constitutivos de seu conteúdo, que expressam as atividades e a relação estabelecida com o seu órgão produtor. A caracterização dos atributos que apresentam o órgão produtor, suas funções e atividades, e a maneira como estes vínculos estabelecidos se traduzem nas partes do

¹⁰³ Destacam-se as propostas metodológicas desenvolvidas pelos Grupos de Identificação de Fundos Internos e o de Fundos Externos (GIFI e GIFE) para solucionar as questões de transferências e recolhimentos de fundos, assim como a metodologia de levantamento da produção documental desenvolvida pela Divisão de Pré-Arquivo para fins de avaliação de documentos no Ministério da Agricultura. (RODRIGUES, 2008).

documento arquivístico, integram o conhecimento gerado nesta fase e compõem a base teórica da identificação arquivística.

Esta pesquisa pode ser desenvolvida durante todas as fases do ciclo de vida dos documentos, podendo, portanto, incidir sobre o momento de sua produção, para efeito de implantação de programas de gestão de documentos, ou no momento de sua acumulação, para controlar fundos transferidos ou recolhidos aos arquivos.

No Brasil, o conceito de identificação remete também ao controle físico dos arquivos, propondo a coleta de dados e registro de informações sobre a situação de acumulação dos documentos. (RODRIGUES, 2012)

No contexto da identificação, a etapa da identificação de tipologias documentais encontra na abordagem da diplomática contemporânea seus fundamentos teóricos e metodológicos, demonstrando a efetiva contribuição desta disciplina para a construção teórica da arquivística.

Nesta perspectiva, a identificação arquivística pode ser utilizada como metodologia de pesquisa para gerar os requisitos de gestão de documentos, de arquivos e de acesso à informação, fundamentando o desenvolvimento desta pesquisa.

4 CONSIDERAÇÕES FINAIS

O tema da identificação, particularmente o da identificação de documentos fundamentada nos parâmetros da diplomática para realizar tarefas padronizadas de classificação, avaliação e descrição, reflete preocupações de pesquisa exploradas pela autora desde 2003, buscando a discussão científica em torno da *identificação* aplicada a situações arquivísticas de produção e acumulação de documentos.

A Espanha tem dado uma efetiva contribuição para a construção da metodologia da identificação no campo da arquivística, produzindo estudos teóricos e aplicados que destacou o país no cenário internacional. O país avançou muito nestes últimos anos na aplicabilidade de práticas normalizadas de identificação de tipos documentais, associando-as a padrões de gestão de documentos convencionais e eletrônicos e, também, na reflexão teórica sobre o assunto. Apenas para exemplificar o avanço nas discussões teóricas sobre o assunto da identificação, que permeia as preocupações dos teóricos espanhóis, o tema da identificação de tipos documentais vem sendo tratado por vários grupos de trabalhos, entre os quais se destacam o Grupo de Arquivistas Municipais de Madri, pioneiro em aplicar a metodologia para a identificação de tipologias de documentos públicos para a realização de tarefas arquivísticas, os Grupos de Arquivistas do Ministério de Cultura da Espanha, do Governo da

Catalunha, dos arquivistas que integram a Mesa de Trabalho de Arquivos da Administração Local e o Grupo de Arquivistas de Universidades espanholas. Os resultados alcançados nestas experiências demonstram a efetiva contribuição desta ferramenta metodológica para a melhoria do funcionamento de sistemas administrativos no contexto da gestão documental na administração pública espanhola e justificam o desenvolvimento parcial deste projeto naquele país.

A padronização de procedimentos de gestão de documentos e de tratamento técnico de arquivos vem sendo tratadas no âmbito de pesquisas desenvolvidas em nível de mestrado¹⁰⁴, cujos procedimentos foram aplicados no estudo de caso desenvolvido no âmbito de administração pública municipal e de doutorado¹⁰⁵, pesquisa que teve como objetivo analisar os processos de identificação de documentos, abordando os aspectos teóricos e metodológicos da tipologia documental e as influências recebidas da diplomática como fundamento para o desenvolvimento da identificação arquivística.

Como desdobramento destas pesquisas, verificou-se a necessidade de aprofundamento do tema, visando estabelecer diálogos entre a identificação, a classificação e a avaliação no contexto da gestão documental, como contribuição para a definição de padrões metodológicos que subsidiem a aplicação da Lei de Acesso à Informação, perspectiva tratada nesta pesquisa de pós-doutorado¹⁰⁶.

O desafio de implantar políticas públicas para a gestão de documentos e arquivos pode ser superado, se pautado em estratégias de cooperação, sobretudo institucionais. Nesta perspectiva, a universidade deve unir-se à sociedade para debater questões novas, que exigem esforços interdisciplinares entre os profissionais envolvidos com a produção e tratamento de documentos e informações, em busca de referenciais teóricos e metodológicos que contribuam para o aperfeiçoamento da administração pública.

Pesquisa coordenada pela autora no âmbito do Programa de Gestão de Documentos do Governo do Estado do Rio de Janeiro¹⁰⁷, RJ demonstram a pertinência da aplicabilidade da

¹⁰⁴ RODRIGUES, 2003.

¹⁰⁵ RODRIGUES, 2008.

¹⁰⁶ Este projeto de pesquisa de pós doutorado “Identificação arquivística como requisito da gestão de documentos: contribuição para a aplicação da Lei de Acesso à Informação no Brasil”, esta sendo desenvolvido parcialmente na Universidad Carlos III de Madrid, Espanha e na Universidade de São Paulo.

¹⁰⁷ Projeto de pesquisa “Identificação arquivística: utilizando a diplomática contemporânea como fundamento metodológico no Programa de Gestão de Documentos do Governo do Rio de Janeiro (PGD-RJ)”, financiado pelo Programa Jovem Pesquisador UFF 2009 (PROPPi/UFF), desenvolvido através de parceria institucional entre o Arquivo Público do Estado / Casa Civil do

metodologia da identificação como requisito da gestão e acesso aos documentos públicos, permitindo a padronização de procedimentos para reconhecer e tratar documentos e informações exigidas para a aplicação da LAI. O estabelecimento desta parceria institucional entre a universidade e governo, demonstra a responsabilidade social da universidade, repassando para a sociedade conhecimentos por ela produzidos que propiciem a melhoria de procedimentos a serem adotados no tratamento técnico dos documentos e acesso às informações do patrimônio público documental.

Este projeto encontra-se em desenvolvimento, cujos resultados preliminares foram apresentados no seminário “Identificação, gestão de documentos y acesso à informação¹⁰⁸”, realizado em maio de 2014 na Universidad Carlos III de Madrid e discutidos nas mesas de debates: “Identificação: termo e conceito no contexto da arquivística” e “Identificação e gestão de documentos: perspectivas metodológicas”, nas quais estiveram presentes Maria Luiza Conde Villaverde (1992), José Luis La Torre Merino (2000) e Concepción Mendo Carmona (2004) autores que fundamentam o referencial teórico desta pesquisa e Julia Maria Rodriguez Barredo, membro do Grupo de Arquivistas Municipais de Madri, que desenvolveu e aplicou o modelo de identificação de tipologia documental publicado por Vicenta Cortes Alonso (1986).

Com a vigência da Lei de Acesso à Informação, a demanda social por documentos e informações que caracteriza o Estado Brasileiro, se amplia cada vez mais. A LAI coloca o Brasil no cenário dos 91 países, entre eles 15 da América Latina, que reconhecem o acesso amplo a documentos e informação.

Governo do Estado do Rio de Janeiro e o Programa de Pós-Graduação em Ciência da Informação da Universidade Federal Fluminense, sob coordenação da autora. Os instrumentos encontram-se disponíveis em: <http://www.aperj.rj.gov.br/instrumentos.htm>.

¹⁰⁸ Seminário "Identificação, gestão de documentos e acesso à informação", organizado pelo Dept ° de Biblioteca e Documentação, através do Programa de Doutorado em Documentação e do Mestrado em Arquivos, Gestão de Documentos e Continuidade Digital da Universidad Carlos III de Madrid, Espanha, no dia 13 de maio de 2014. Os vídeos estão disponíveis em: 1. Abertura e apresentação do projeto de pesquisa “Identificação arquivística como requisito da gestão de documentos: contribuição para aplicação da Lei de Acesso à Informação no Brasil”, apoio CAPES, <https://arcamm.uc3m.es/arcamm/?item=e12764a0ac25abb85580213af81f880b>; Mesa 1 “Identificação: termo e conceito no contexto da arquivística” <https://arcamm.uc3m.es/arcamm/?item=2908d624614240c3fd1d1b6c1cb584c6> e Mesa 2 “Identificação e gestão de documentos: perspectivas metodológicas” <https://arcamm.uc3m.es/arcamm/?item=2c23070489f58bf3dad1b891540e09f0>. Acesso em: 29 jul 2014.

Os documentos públicos são peças que estão integradas num todo, orgânico e estruturado, que lhes conferem sentido de existir, ou seja, a razão de ser da própria administração. Significa dizer que o documento só preserva seu valor quando está integrado num conjunto próprio, que reflete o contexto de sua produção, que expressa o valor probatório sobre o funcionamento do governo, do exercício dos direitos e deveres do cidadão.

A identificação arquivística é uma metodologia de pesquisa que se fundamenta em estudos sobre o documento de arquivo e seu órgão produtor, decorrente de diagnóstico elaborado, com a finalidade de propor soluções para o problema apresentado.

Os estudos sobre o documento arquivístico e o vínculo que mantém com o contexto administrativo em que é produzido, perspectiva da identificação arquivística, vêm nutrindo o debate científico internacional sobre a formulação de requisitos para a gestão documental.

Trata-se de uma questão emergente no âmbito do debate científico, em busca de soluções para superar os desafios impostos para a plena aplicabilidade da Lei de Acesso à Informação pública no Brasil.

REFERÊNCIAS

ABRUCIO, Fernando. **A Lei de Acesso à Informação e a cidadania**. Revista Época. Disponível em: <http://revistaepoca.globo.com/opiniaofernando-abrucio/noticia/2012/06/lei-de-acesso-informacao-e-cidadania.html>. Acesso em: 08 ago 2012.

ARQUIVO NACIONAL. Gestão de documentos; conceitos e procedimentos básicos. Rio de Janeiro, 1995. (Publicações Técnicas, 47)

ARQUIVO PÚBLICO DO ESTADO DE SÃO PAULO (APESP). **Lei de Acesso a Informação**. Publicado em 31 de maio de 2012. Disponível em: http://www.arquivoestado.sp.gov.br/noticia_ver.php?id=338. Acesso em: 08 ago 2012.

BAUDRIT BARQUERO, Álvaro. El derecho a la información. In: Actas del Seminario Internacional de Normatividad Archivística, 1999. **Actas...** Lima (Peru): Archivo General de la Nación, 1999. (cd)

BELLOTTO, Heloisa Liberalli. Identificação de fundos. In: **Arquivos permanentes; tratamento documental**. 2. ed. Rio de Janeiro: FGV Editora, 2004. p.127-134.

BERTAZZI, Danilo Marasca. **O projeto de lei de acesso à informação e seu impacto sobre os servidores públicos**. Artigo 19. Estudos em Liberdade de Informação, 2011. Disponível em: <http://gregmichener.com/Leis-de-Acesso-a-Informacao-Dilemas-da-Implementacao-Artigo19.pdf> Acesso em: 17 fev. 2013. (série Estudos em Liberdade de

BRASIL. **Constituição** da República Federativa do Brasil. Brasília: Senado **Federal**, 1988.

BRASIL. **Lei n. 12.527**, de 18 de novembro de 2011. Lei de Acesso à Informação. Disponível em <http://www.acaoainformacao.gov.br/acaoainformacaogov/acao-informacao-brasil/legislacao-integra-completa.asp>. Acesso em 08 ago 2012.

BRASIL. **Lei n. 8.159**, de 8 de janeiro de 1991. Dispõe sobre a política nacional de arquivos públicos e privados e dá outras providências. Diário Oficial [da] República Federativa do Brasil, Brasília, DF, 9 jan. 1991. Seção 1, p. 457.

CARMONA MENDO, Concepción. Consideraciones sobre el metodo en archivística. In: **Documenta & Instrumenta**. Madrid, Universidad Complutense de Madrid, v. I, 2004, p. 35-46. Disponível em <http://www.ucm.es/info/document/htm/default.htm>, acesso em 31 de agosto de 2005.

CGU. **Acesso à Informação Pública**: uma introdução à Lei 12.527, de 18 de novembro de 2011. Cartilha da Controladoria-Geral da União. Brasília, Imprensa Nacional 2011. Disponível em: www.acs.ufpr.br/cgu.pdf. Acesso em: 08 ago 2012.

CONDE VILLAVERDE, Maria Luisa. **Manual de tratamiento de archivos administrativos**. Madrid: Dirección de Archivos Estatales / Ministério de Cultura, 1992.

CORTÉS ALONSO, Vicenta. Nuestro modelo de análisis documental. Boletín de ANABAD (Asociación Española de Archiveros, Bibliotecarios, Museólogos y Documentalistas). Madrid: **ANABAD**, n. 3, p. 419-434, 1986.

DA MATTA, Roberto. **Sumário Executivo** - Diagnóstico sobre Valores, Conhecimento e Cultura de Acesso à Informação Pública no Poder Executivo Federal Brasileiro. Disponível em: http://www.acessoainformacao.gov.br/acessoainformacaogov/publicacoes/SUMARIO_FINAL.pdf. Acesso em: 17 fev. 2013.

DICCIONARIO de terminología archivística. Madrid: Dirección de Archivos Estatales / Ministerio de Cultura, 1993.

DICIONÁRIO brasileiro de terminologia arquivística. Rio de Janeiro: Arquivo Nacional, 2005.

DICIONÁRIO de terminologia arquivística. Ana Maria de Almeida Camargo, Heloísa Liberalli Bellotto (coord.). São Paulo: Associação dos Arquivistas Brasileiros - Núcleo Regional de São Paulo: Secretaria de Estado da Cultura, 1996.

DURANTI, Luciana. **Diplomatica**: nuevos usos para una antigua ciencia. Trad. Manuel Vazquez. Carmona: Asociación de Archiveros de Andalucía, 1995.

JARDIM, José Maria. A Lei de Acesso à Informação Pública: dimensões político-informacionais. XIII ENANCIB. **Anais...** Rio de Janeiro, 2012. Disponível em: <http://obgi.files.wordpress.com/2012/10/jardim-xiii-enancib.pdf>

LA TORRE MERINO, José Luis & MARTÍN-PALOMINO y BENITO, Mercedes. **Metodología para la identificación y valoración de fondos documentales**. Madrid: Ministerio de Educación, Cultura y Deportes. S.G. de Información y Publicaciones, 2000. (Escuela Iberoamericana de Archivos: experiencias y materiales)

LÓPEZ GÓMEZ, Pedro. Los archiveros y sus investigaciones. **Métodos de Información**. v. 5, n. 22-23, 1998, p. 37-43. Disponível em <http://eprints.rclis.org/archive/00001743/>, acesso em 20 de setembro de 2005.

MICHENER, Greg. **Liberdade de Informação**: uma Síntese dos Dilemas de Conformidade suas Possíveis Soluções. Artigo 19. Estudos em Liberdade de Informação, 2011 Disponível

em: <http://gregmichener.com/Leis-de-Acesso-a-Informacao-Dilemas-da-Implementacao-Artigo19.pdf> Acesso em: 17 fev 2013.

PRATAS, Sérgio. **O acesso à informação administrativa no século XXI**. Disponível em: http://www.cada.pt/uploads/estudos/O_Acesso_Informacao_Administrativa_no_seculo_XXI.pdf. Acesso em: 14 ago 2012.

RODRIGUES, Ana Célia. **Diplomática contemporânea como fundamento metodológico da identificação de tipologia documental em arquivos**. São Paulo: Universidade de São Paulo, 2008. Tese (Doutorado em História Social). (Mestrado em História Social). Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, SP. Disponível em: www.teses.usp.br/teses/.../8/.../TESE_ANA_CELIA_RODRIGUES.pdf. Acesso em: 26 set. 2010.

RODRIGUES, Ana Célia. **Identificação: uma metodologia de pesquisa para a Arquivística**. In: Estudos avançados em Arquivologia. Marília, São Paulo: Oficina Universitária; Cultura Acadêmica, 2012, p. 197-215. Disponível em: http://www.marilia.unesp.br/Home/Publicacoes/estudos_avancados_arquivologia.pdf. Acesso em: 20 mar. 2014.

RODRIGUES, Ana Célia. **Tipologia documental como parâmetro para a gestão de documentos de arquivos: um manual para o Município de Campo Belo, MG**. São Paulo: Universidade de São Paulo, 2003. (Dissertação de Mestrado). Disponível em: www.teses.usp.br. Acesso em: 26 jun 2007.

VIEIRA, Vânia Lúcia Ribeiro. **A Lei de Acesso à Informação**. 2012. Disponível em: http://bdjur.stj.jus.br/xmlui/bitstream/handle/2011/46754/Acesso_informacao_texto.pdf?sequence=1. Acesso em: 17 fev 2013.

ORGANIZAÇÃO E REPRESENTAÇÃO DA INFORMAÇÃO NO MUSEU VIRTUAL DE INSTRUMENTOS MÚSICAIS – MVIM

ORGANIZATION AND REPRESENTATION OF INFORMATION IN THE VIRTUAL MUSEUM OF MUSICAL INSTRUMENTS - MVIM

Adriana Olinto Ballesté
Alea de Almeida

Resumo: O *Museu Virtual de Instrumentos Musicais-MVIM* é um projeto proposto pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) em parceria com a Escola de Música da Universidade Federal do Rio de Janeiro (EM-UFRJ) e financiamento da Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ), que visa à divulgação de acervos variados e distintos de instrumentos musicais em um ambiente virtual. Nesse artigo, priorizamos a discussão sobre a adaptação, organização e representação do primeiro acervo a ser disponibilizado no MVIM: o acervo do Museu Instrumental Delgado de Carvalho da EM-UFRJ. Mostramos a importância do acervo do Museu Instrumental Delgado de Carvalho e as lacunas em relação às informações pré-existentes. Discutimos as questões metodológicas em relação à organização e representação da informação na Musicologia e na Organização do Conhecimento, procurando mostrar pontos convergentes entre as duas áreas e aqueles que sejam relevantes para o MVIM. Salientamos a problemática em relação à classificação, representação e descrição dos instrumentos do Museu Delgado de Carvalho no catálogo virtual do MVIM pontuando com os padrões, normas, regras e soluções adotadas por museus e instituições internacionais relacionadas às áreas de música e de instrumentos musicais.

Palavras-chave: Organização da Informação. Catalogação. Museu Virtual. Classificação de instrumentos musicais. Recuperação da Informação. Instrumentos Musicais.

Abstract: The Virtual Museum of Musical Instruments-MVIM is proposed by the Brazilian Institute for Information in Science and Technology (IBICT) in partnership with the School of Music of the Federal University of Rio de Janeiro (EM-UFRJ) and funding of Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ), which aims to disseminate varied and distinctive collections of musical instruments in a virtual environment. In this article, we prioritize the discussion on adaptation, organization and representation of the first collection to be made available in MVIM: the collection of the Museu Instrumental Delgado de Carvalho EM-UFRJ. We show the importance of the Museu Instrumental Delgado de Carvalho and gaps in relation to pre-existing information. We discuss methodological issues regarding the organization and representation of information in Musicology and Knowledge Organization, aiming to show convergence between the two areas and those that are relevant to the MVIM. We underline the issue regarding the classification, representation and description of the Museu Delgado de Carvalho in the virtual catalog of MVIM scoring with standards, norms, and rules adopted by international museums and institutions related to the areas of music and musical instruments solutions.

Keywords: Information Organization. Cataloging. Virtual Museum. Classification of musical instruments. Information Retrieval. Musical instruments.

1 INTRODUÇÃO

O *Museu Virtual de Instrumentos Musicais – MVIM*, projeto proposto pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) em parceria com a Escola de

Música da Universidade Federal do Rio de Janeiro (EM-UFRJ) e financiamento da Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ) ¹⁰⁹, visa à divulgação de acervos variados e distintos de instrumentos musicais – que podem ser de outros museus, instituições ou pessoas físicas – em um espaço virtual.

No MVIM os acervos poderão ser vistos em conjunto e as informações disponibilizadas serão fontes para pesquisas e conexões diversas entre os itens de acervos que no mundo concreto podem ou não estar em localizações distintas.

O primeiro acervo a ser disponibilizado no MVIM é o do Museu Instrumental Delgado de Carvalho, vinculado à Escola de Música da UFRJ (CARDOSO, 2008) ¹¹⁰. O acervo deste museu é composto por aproximadamente 80 instrumentos musicais de épocas e origens diversas e, desde 2008, está indisponível para o público em geral. Além de estar sendo organizado para estar disponível de forma virtual na Web, o acervo do Museu da UFRJ está sendo higienizado, restaurado e acondicionado, fotografado, catalogado e organizado fisicamente.

Na primeira versão do MVIM o ponto central é o “catálogo virtual” dos instrumentos musicais, onde podem ser encontrados os seguintes grupos de informação: identificação e dados gerais do instrumento; dados específicos do exemplar; informações históricas e curiosidades do instrumento; registros audiovisuais; referências bibliográficas. No espaço virtual do museu estão presentes também as seções: “em pauta”, um espaço livre em permanente renovação, com opiniões exposições e curiosidades sobre música e instrumentos musicais; “artigos, entrevistas” um espaço para artigos acadêmicos; “normas para a elaboração do catálogo”; “referências bibliográficas”; informações gerais do “projeto” e; links para sítios relacionados à música e instrumentos musicais. Para a segunda versão do MVIM estamos preparando um espaço educativo dinâmico que visa a divulgação de acervos musicais de forma a atingir um público variado, mas com foco principal nos estudantes e professores do Ensino Fundamental.

Nesse artigo vamos apresentar e discutir questões relacionadas à organização de informações em um museu virtual de instrumentos musicais - classificação, padronização, descrição dos itens do acervo - fazendo um contraponto entre metodologias, regras e normas internacionais. Iniciamos o artigo contextualizando a trajetória e a importância do primeiro

¹⁰⁹ O Museu Virtual de Instrumentos Musicais estará disponível ao público na web a partir de setembro de 2014.

¹¹⁰ Artigo sobre o Museu disponível em: <http://www.musica.ufrj.br/index.php?option=com_content&view=article&id=79&Itemid=121>. Acesso em: 25 abr.2011.

museu de instrumentos musicais do Brasil, o Museu Instrumental Delgado de Carvalho e a lacuna e problemáticas dos inventários e informações preexistentes. Discutimos as questões metodológicas relacionadas à organização e representação da informação na Musicologia e na Organização do Conhecimento, procurando mostrar pontos convergentes entre as duas áreas e aqueles que sejam relevantes para o MVIM. Ressaltamos a problemática em relação à classificação, representação e descrição dos instrumentos do Museu Delgado de Carvalho no catálogo virtual do MVIM. E finalmente, fazemos considerações sobre possibilidades de melhorias para a área de organização e terminologia de acervos de instrumentos musicais.

2 MUSEU INSTRUMENTAL DELGADO DE CARVALHO

A música no Brasil sempre foi muito valorizada. Já nos primeiros relatos escritos pelos navegantes portugueses é marcante a menção à música e aos instrumentos musicais. Fernão Cardim (1978), em cartas durante sua viagem ao Brasil entre os anos de 1583 e 1590, conta que os índios recebiam os visitantes "uns cantando e tangendo a seu modo [...] dançando ao som da viola, pandeiro, tamboril e flauta [...] os padres a ler, a contar, a cantar e tanger [...]".

Com a chegada da Corte Portuguesa, em 1808, os ares das grandes cidades européias são sentidos no Brasil. O Rio de Janeiro, cidade sede da Corte, se transforma no centro cultural mais importante do país. Durante todo o século XIX, se mantém esse aquecimento na área cultural.

Sob a influência desses novos ares e hábitos culturais, no final do século XIX, é criado o primeiro museu de instrumentos musicais no Brasil no Instituto Nacional de Música¹¹¹, pelo seu primeiro diretor, o compositor e maestro Leopoldo Miguéz (1850 a 1902). O Museu destinava-se, segundo o Regulamento do Instituto Nacional de Música de 1900, citado por Brandão (2013, p. 42), ao estudo de história de música e organologia musical.

O museu tinha em seu acervo instrumentos provenientes de diversos países, como Java, Egito, Sudão, Índia, China, Estados Unidos da América, Marrocos e Pérsia (Irã). O seu acesso era restrito, sendo o ingresso na sala do museu permitido somente aos alunos acompanhados por professores e com a autorização do diretor.

A partir da década de 1970, o museu é aberto ao público ficando seu acervo exposto em vitrines no corredor principal da Escola de Música até 2008, quando o museu foi

¹¹¹ O Instituto Nacional de Música foi criado após a Proclamação da República, em 1889, derivado do Conservatório de Música, criado em 1848 no Rio de Janeiro, por Francisco Manoel da Silva (1795-1865). Em 1937, a Universidade do Rio de Janeiro encampa o Instituto e esse passa a se chamar Escola Nacional de Música, onde hoje está abrigado o Museu Delgado de Carvalho.

desativado e os itens armazenados na Biblioteca Alberto Nepomuceno da Escola de Música (CARDOSO, 2008).

3 OS INVENTÁRIOS DO MUSEU INSTRUMENTAL DELGADO DE CARVALHO

Desde a sua criação foram gerados inventários dos itens constitutivos do Museu Instrumental Delgado de Carvalho. O mais antigo, um inventário manuscrito por Leopoldo Miguéz, entre os anos de 1890 e 1895, descrevia os primeiros itens do acervo, e foi redescoberto após o início do projeto. Nele constam 49 instrumentos de diversas nacionalidades – Síria, Índia, Marrocos, Sudão, China, México, Estados Unidos e Brasil –, porém desses apenas 27 estão presentes atualmente no acervo.

Em 1905, foi publicado um inventário, organizado e classificado pelo compositor Joaquim Tomas Delgado de Carvalho (1872-1922), que assumiu a responsabilidade, como ele mesmo afirma, pela “inspeccção do Museu instrumental, Gabinete de acústica e Bibliotheca” (CARVALHO, 1905, p. 5). Esse inventário, mais descritivo, mostra que o acervo cresceu bastante, contando, então, com 87 instrumentos musicais, dentre os tradicionais instrumentos de orquestra – violinos, violas, fagotes, oboé, bandolins napolitanos – até instrumentos de culturas variadas e distantes como um *gi-hin*, instrumento de cordas chinês; uma *darabuka*, tambor de origem egípcia e um *dog-dog*, tambor de bambu de Java. Além dos instrumentos o museu possuía 54 itens diversos, como cartas, bilhetes, cartões postais, autógrafos.

Somente meio século depois, na década de 1970, foram elaborados novos inventários, sendo alguns descobertos recentemente na biblioteca da Escola de Música, quando então o museu é aberto ao público, ficando seu acervo exposto em vitrines no corredor principal da Escola de Música. Em 2008, com a desativação do museu, foi feita uma relação dos itens e a partir de 2011 teve início um novo inventário, finalizado em 2013 (BRANDÃO, 2013) com o auxílio de professores de música¹¹². Depois disso, os itens foram catalogados na Base Minerva da UFRJ. A lista completa de inventários pode ser vista no

¹¹² A identificação dos instrumentos contou com a colaboração de diversos professores da Escola de Música da UFRJ e de outras universidades: Alysio de Mattos, Paulo Sá e Márcia Taborda (instrumentos de cordas); Eduardo Monteiro, José Rua, Cristiano Alves e David Alves (sopros); Pedro Sá e Samuel Araújo (percussão) e ainda professores de outras universidades: Rogério Budaz (Universidade da Califórnia); Patrícia Aguillar e Mônica Lucas (Universidade de São Paulo).

QUADRO 3.

QUADRO 3 – Inventários do Museu Instrumental Delgado de Carvalho

1890-1895	Livro de Inventário de Leopoldo Miguéz
1905	Catálogo organizado por Delgado de Carvalho e publicado pela Imprensa Nacional.
1973	Anotações manuscritas feitas por Mary Hugo Braga Pinto Coelho ao catálogo de 1905.
1974	Inventário datilografado organizado por Luciano Rolla, com base registros de 1905 e de 1973.
1990	Inventário organizado e atualizado por Léo Soares com base nos inventários anteriores.
1994	Inventário organizado por Afifi Craveiro de Almeida publicado na Revista Brasileira de Música intitulado “O Museu Instrumental Delgado de Carvalho: Breve Notícia”
2008	Relação organizada por Dolores Castorino Brandão, bibliotecária chefe da Biblioteca Alberto Nepomuceno da Escola de Música da UFRJ.
2013	Inventário organizado por Dolores Castorino Brandão.

Fonte: as autoras.

Esses inventários mostram sem dúvida a evolução do acervo, mas apresentam lacunas e imprecisões de nomenclatura, de classificação, de descrição e outros problemas. O contato com esses documentos, o estudo de regras e normas adotadas em outros museus de música e o estudo da literatura especializada da área de Organização do Conhecimento e da Musicologia nos mostraram a importância e a necessidade de se ter critérios bastante rigorosos na descrição de um acervo.

4 ORGANIZAÇÃO DO CONHECIMENTO E A MUSICOLOGIA

No final do século XX, o desenvolvimento tecnológico estimula a criação, organização, desenvolvimento e disponibilização de catálogos, coleções e bibliotecas digitais relacionadas à música, incentivando os estudos interdisciplinares. Lanzelotte, Ulhoa e Ballesté (2004) evidenciam que trabalhos que envolvem a disponibilização na web de acervos de

música abrangem dentre outros aspectos, diversas formas de organização e representação da informação musical que por sua complexidade remetem à necessidade de estudos interdisciplinares entre a Musicologia e outras áreas do conhecimento como a Organização do Conhecimento.

O registro, a classificação, a representação, a organização do conhecimento “tem sido preocupação do homem que remonta aos tempos de sua própria existência” (MIRANDA, 1999, p. 64). Desde o século III a.C. tanto na Organização do Conhecimento quanto na Musicologia já havia uma preocupação com a sistematização do conhecimento. Aristóteles (384 – 322 a.C.), na Grécia, é considerado o primeiro teórico sistemático do conhecimento, pois é evidente em suas idéias uma grande preocupação em saber como o homem conhece e quais as condições do conhecimento verdadeiro (BURKE, 2000, p. 19). Na China, nesse mesmo século, foi elaborado o mais antigo sistema de classificação de instrumentos musicais que se tem notícia, que os classifica de acordo com seus materiais constitutivos: pedra, pele, madeira, metal, bambu, seda, couro. (KARTOMI, 1990).

Desde então vários sistemas de classificação e organização do conhecimento foram criados e/ou recriados. Hoje estão ainda em uso sistemas do século XIX, quando a atividade de classificação e organização do conhecimento ganha uma importância muito grande. Os primeiros sistemas de classificação usavam uma *abordagem universal* que procurava englobar todos os assuntos ou disciplinas e os organizava, em geral, em um esquema enumerativo. Alguns desses têm hoje ainda uma "utilidade indiscutível, quando se trata de localizar fisicamente o documento em sua classe de assunto" (DODEBEI, 2002).

Na Biblioteconomia um relevante exemplo desse tipo de sistema é o *Dewey Decimal Classification - CDD*, que é organizado numericamente de forma decimal e hierarquicamente com base em disciplinas e subdisciplinas divididas em dez classes principais que são subdivididas em outras dez e assim sucessivamente. As Artes são representadas na classe 700 e a Música na classe 780 e suas principais disciplinas, representadas conforme o QUADRO 4.

QUADRO 4 – Classes principais representativas da música

- 780 Música
- 781 Princípios gerais e formas musicais
- 782 Música vocal
- 783 Música para uma só voz
- 784 Instrumentos & Conjuntos instrumentais
- 785 Música de câmara
- 786 Teclado & outros instrumentos
- 787 Instrumentos de corda (Cordofones)

788 Instrumentos de sopro (Aerofones)

Fonte: Dewey Decimal Classification (tradução das autoras).

A Classificação Decimal Dewey pretende ser muito abrangente, e por isso têm restrições que não permitem que o sistema de conta de particularidades. Mas não podemos negar a importância do CDD, que foi um marco para a biblioteconomia, deixou sementes para a concepção de outros sistemas de classificação e ainda hoje é um dos mais utilizados em bibliotecas no mundo todo.

Na Musicologia o sistema de classificação de instrumentos musicais mais utilizado, concebido por Curt Sachs e Eric Von Hornbostel, em 1914, toma como base para sua organização numérica o sistema *CDD*. Os instrumentos, nesse sistema, são divididos em quatro grupos, de acordo com a forma de produção do som, ou seja, com a vibração do instrumento: os (1) *idiofones*, instrumentos rígidos que produzem o som pela vibração do seu próprio corpo, como as baquetas, os pratos e triângulos; os (2) *membranofones*, ou instrumentos de membranas, nos quais o som é produzido pela contração e descontração de uma membrana, como o tambor; os (3) *cordofones*, ou instrumentos de cordas, em que a vibração das cordas produz o som como o violão e o violino; e os (4) *aerofones*, ou instrumentos de “ar”, em que uma coluna de ar vibra produzindo o som, como a flauta e a trompa.

Sachs (1968), um dos autores desse sistema de classificação, afirma que existem imperfeições nesse sistema, bem como em qualquer sistema de classificação de instrumentos que pretenda ser generalizante, devido a uma lógica de existência “artificial”.

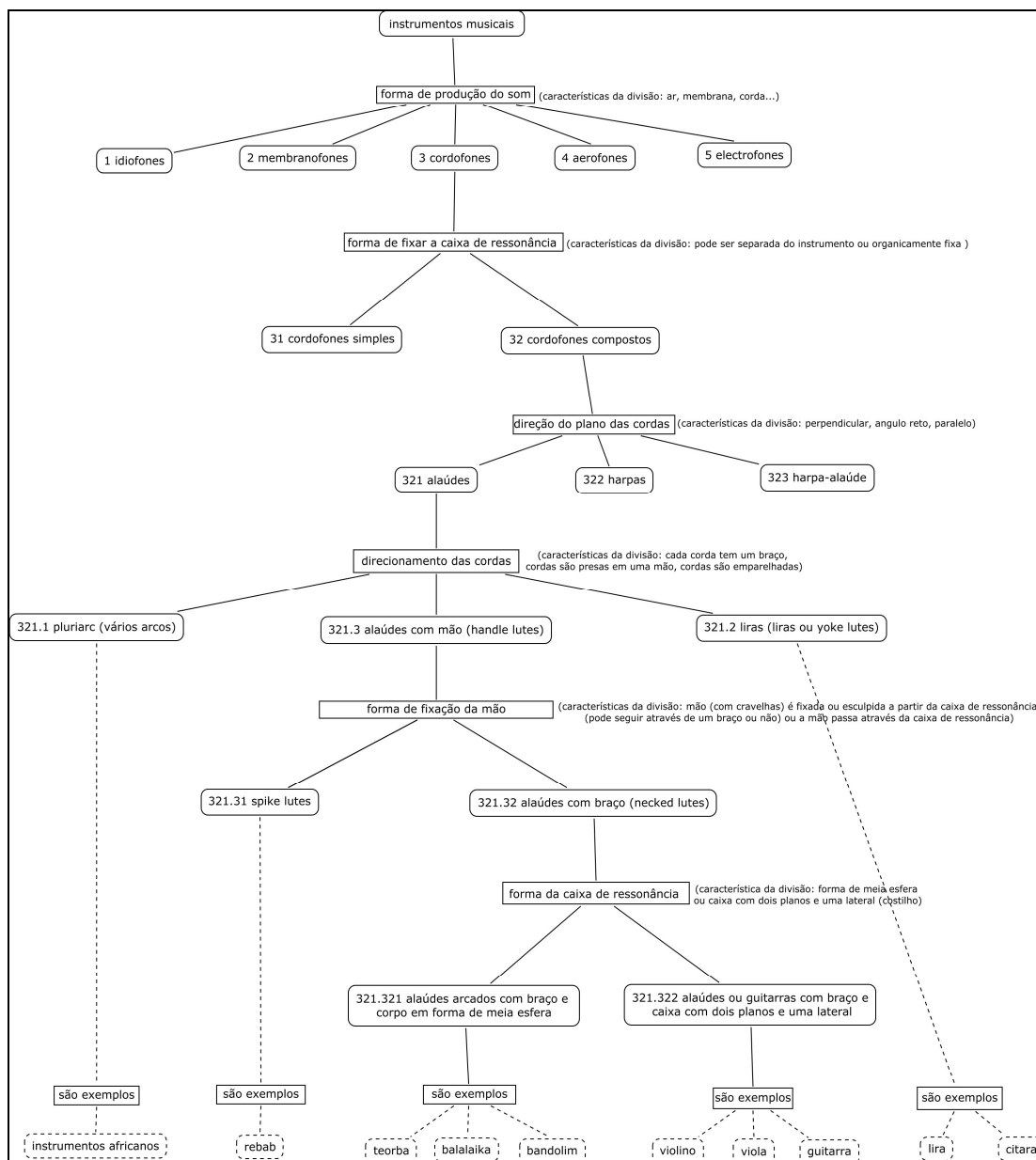
Depois do sistema de Hornbostel&Sachs outras soluções foram propostas para classificação de instrumentos musicais. André Schaeffner (1936) propõe uma divisão principal para os instrumentos, não em quatro, mas em duas classes: instrumentos de corpos sólidos e instrumentos que vibram com o ar. Hans-Heinz Dräger, em meados do século XX, criou uma classe, chamada *electrofone* que se baseia em vibrações elétricas. Olsen, em 1980, propõe a incorporação do corpo humano como mais uma categoria de instrumentos, com o nome de *corpophones*. (KARTOMI, 2001, p. 286).

A partir dessas propostas foram criados apêndices, adendos e reduções, mas nada que fugisse à base original do sistema de Hornbostel&Sachs, que até hoje é o mais utilizado nos museus e coleções de instrumentos musicais.

Na FIGURA 7 representa-se a classificação de Hornbostel & Sachs, usando um mapa conceitual¹¹³, incluindo a nova classe adotada há pouco tempo, os (5) *electrofonos*. Nessa figura, visando uma simplificação, foi privilegiado o ramo da classificação dos cordofones. Os cordofones se subdividem de acordo com a forma de fixação da caixa de ressonância em: (31) cordofones simples e (32) cordofones compostos. Esses últimos se dividem de acordo com a direção do plano das cordas em: (321) alaúdes, (322) harpas, (323) harpa-alaúde. E assim sucessivamente vão sendo geradas novas divisões.

FIGURA 7 – Representação da classificação de Hornbostel & Sachs (1914), utilizando um mapa conceitual, privilegiando a exibição do ramo de cordofones.

¹¹³ *Mapas Conceituais* são ferramentas gráficas que visam à organização e representação do conhecimento. São estruturados a partir de conceitos fundamentais e suas relações. Normalmente os conceitos são representados em caixas de texto e a relação entre os conceitos em uma linha de conexão.



Fonte: Ballesté (2009).

Em meados do século XX os sistemas de classificação enumerativos que procuram dar conta de uma classificação generalista perdem espaço para sistemas de classificação analítico-sintéticos, com uma abordagem contextual, que possibilita a inclusão de categorias de classificação a posteriori e de acordo com a área examinada. Dois marcos teóricos importantes foram a Teoria da Classificação Facetada e a Teoria Geral da Terminologia (TGT). A primeira foi concebida por Shiyali Ramamrita Ranganathan, na década de 1930, e estabelece novos princípios para a classificação bibliográfica visando tanto à organização dos livros em estantes quanto à organização do conteúdo dos livros. Dessa forma, o mapeamento de um domínio do conhecimento poderia ser feito de forma independente e as classes poderiam ser criadas “no momento em que o livro era examinado de acordo com os elementos

conceituais do seu assunto, e sintetizado de acordo com regras das fórmulas facetadas dos limites das disciplinas” (DAHLBERG, 1976, p. 86).

A Teoria Geral da Terminologia foi desenvolvida por Eugen Würster, em 1931, com o objetivo de uniformizar conceitos e termos por meio de vocabulários controlados, garantindo a comunicação precisa entre profissionais de um mesmo universo do conhecimento, sendo uma de suas grandes contribuições a fixação de um espaço de atuação para a Terminologia, diferenciando-a da Lexicografia.

O grande avanço da TGT foi sistematizar princípios terminológicos que deram a esta atividade um caráter científico próprio, diferenciando-a da atividade lexicográfica. Para muitos profissionais envolvidos com as questões de denominação, até aquele momento não estava bem determinado o campo de atuação de cada uma das atividades (CAMPOS, 2001)

Na Musicologia, a Terminologia e a Lexicografia, que são consideradas disciplinas inerentes, ainda se confundem. No dicionário Grove, o mais conceituado na área da Música, por exemplo, as duas são definidas no mesmo verbete, como uma forma de pesquisa acadêmica direcionada a condensar, organizar e explicitar os termos que os músicos usam para comunicar suas idéias sobre suas experiências na arte (DUCKLES & BALCHIN, 2009). De fato essas duas disciplinas têm pontos em comum, mas têm claras distinções. A Lexicografia se concentra no estudo da palavra e todos os seus possíveis significados, uma de suas aplicações é a elaboração de dicionários. A Terminologia estuda o conjunto de termos e conceitos concentrando-se nas relações possíveis entre eles, quando inseridos em uma determinada área do conhecimento, entre suas aplicações está a elaboração de tesouros, ontologias, taxonomias e vocabulários controlados. A falta de uma distinção rigorosa entre as duas disciplinas durante tanto tempo provavelmente é o motivo pela confusão entre Terminologia e Lexicografia dentro da Musicologia.

Até meados do século XX, os musicólogos concentram seus estudos principalmente no campo da Lexicografia, grandes contribuições foram dadas especialmente durante o século XVIII quando os primeiros dicionários e enciclopédias relacionados à Música foram elaborados. Porém, no final do século XX, a Terminologia ainda é uma disciplina que recebe pouca atenção no universo da Musicologia. Uma evidência disso ocorreu no *American Musicological Society*, em 1991, um encontro especial, com o objetivo de tratar das tendências da Musicologia nos anos 1980 e que não contou com nenhum artigo ou referência sobre Terminologia (HOLOMAN&PALISCA, 1982).

Na Organização do Conhecimento, ao contrário, a Terminologia é uma área bastante valorizada, mas ainda assim, segundo Dahlberg (2008), apesar da existência de uma “teoria da

classificação capaz de explicar uma quantidade de fatores anteriormente desconhecidos, ou conhecidos apenas intuitivamente”, é importante dar continuidade às pesquisas na área, pois agora é possível “determinar com exatidão o que era certo ou errado, e saber como os sistemas de classificação podem ser aperfeiçoados” (DAHLBERG, 2008).

No caso dos instrumentos musicais, a inexistência de um sistema terminológico de controle do uso dos termos, como, por exemplo, um tesouro multilíngüe dificulta o trabalho de organização de um acervo e possibilita discrepâncias terminológicas entre distintos catálogos.

5 ORGANIZAÇÃO DA INFORMAÇÃO NO MVIM

Para a organização da informação no Museu Virtual de Instrumentos Musicais – MVIM, inicialmente, estudamos soluções encontradas em outros museus e instituições que lidam com acervos de instrumentos musicais. Percebemos depois a importância e a necessidade de estabelecer um diálogo interdisciplinar entre metodologias da Musicologia e da Organização do Conhecimento. Ficaram evidentes questões tais como: a necessidade de se estabelecer um vocabulário controlado em língua portuguesa; a dificuldade de classificação de alguns instrumentos; a importância de se estabelecer critérios bem definidos e padronizados para a descrição, a determinação da autoria e medidas de cada instrumento; a investigação das marcas e inscrições; a informação sobre a restauração dos instrumentos; a pesquisa e padronização dos registros audiovisuais.

Investigamos inicialmente como estavam organizados os sítios de outros museus de música na Web, tais como: *Museu de Música* de Lisboa; *Musée da Cité de la Musique* de Paris; *Musée des Instruments de Musique* de Bruxelas; *Horniman Public Museum* de Londres; *Musical Instrument Museum Online* (MIMO); *The Metropolitan Museum of Art* de Nova Iorque; e o comitê de museus relacionados a instrumentos musicais, o CIMCIM – *Comité Internacional de Museos y Colecciones de Instrumentos Musicales*¹¹⁴.

Essa investigação, que se tornou permanente, aliada ao acesso a documentos tais como: catálogos impressos; tratados de organologia e classificação de instrumentos; referências e padrões adotados e definidos em conjunto por instituições internacionais; auxiliou-nos a estabelecer e estruturar as normas de catalogação MVIM.

¹¹⁴ CIMCIM - Musical Instruments International Committee for Museums and Collections of Musical Instruments / ICOM – International Council of Museums. Disponível em: <http://icom.museum/>.

O fato de não haver um museu de instrumentos musicais estabelecido no Brasil e a proximidade da cultura e do idioma nos aproximou de instituições portuguesas, por isso uma das fontes que mais influenciou o nosso trabalho foram as “Normas de Inventário dos Instrumentos Musicais” (TRINDADE, 2011) criada para os acervos portugueses.

Com base nessas referências estabelecemos critérios para a estrutura e descrição das informações no sítio do museu e, especialmente, para o catálogo de instrumentos musicais, que, como já mencionamos anteriormente, é o cerne do museu virtual nessa sua primeira versão.

O catálogo de instrumentos engloba os campos, listados no QUADRO 3. A maior parte deles estará disponível no sítio do MVIM e os outros estarão disponíveis apenas para consultas por parte da equipe técnica de desenvolvimento.

QUADRO 5 – Campos do Catálogo do MVIM

Código MVIM	Observações	Restauração
Nome do instrumento	Autoria/fabricante	Notas gerais
Nomes adicionais	Local de fabricação	Bibliografia
Categoria	Data de fabricação	Descrições anteriores
Classificação	Material	Imagens
Localização física	Dimensões	Áudios
Proveniência	Estado de conservação	Vídeos
Descrição	Marcas e inscrições	

Fonte: as autoras.

Para definir a nomenclatura do instrumento utilizamos como base as recomendações do CIMCIM, que tem como missão a promoção de padrões para o uso e a conservação de museus e coleções de instrumentos musicais. O comitê recomenda a utilização da nomenclatura adotada pelo *The New Grove Dictionary of Musical Instruments*, já que não há um sistema para o controle do uso de termos de instrumentos musicais. Para conferir essa nomenclatura padrão utilizamos principalmente o sítio do projeto MIMO - Musical Instrument Museums Online (2013), um consórcio de importantes museus europeus que visa à centralização de acesso on-line às coleções de instrumentos musicais.

Após o acesso a essas fontes passamos a investigar se os nomes dos instrumentos adotados pelos catálogos e inventários anteriores do Museu Instrumental Delgado de Carvalho estavam em consonância com a nomenclatura padrão. Nesse momento percebemos que, apesar do esforço despendido em cada um desses trabalhos precedentes do Museu Instrumental Delgado de Carvalho, existiam incorreções na nomenclatura, na classificação e na descrição dos instrumentos, algumas que se repetem desde o primeiro registro e outras que

surtem nos novos inventários. O fato de não termos os nomes corretos dos instrumentos dificultou a busca no sítio do MIMO, nos dicionários e publicações.

Com o intuito de exemplificar alguns desses problemas mostraremos uma série de procedimentos que fizemos na busca da nomenclatura ou termo padrão para um instrumento do museu, o 'Yueqin', de origem chinesa, que mostramos na FIGURA 8. Esse instrumento possui uma caixa de ressonância arredondada, formada por duas peças circulares de madeira unidas pela costilha. O instrumento possui quatro cordas. O braço é curto e possui duas divisões. Ao longo do tampo harmônico encontram-se mais oito divisões.

FIGURA 8 - Imagem do Yueqin



Iniciamos a pesquisa com o nome 'Yne-Kin' encontrado nos inventários de 1905 e de 1974 do Museu Instrumental Delgado de Carvalho, pois ainda não tínhamos acesso aos outros inventários¹¹⁵. Buscamos no catálogo do Projeto MIMO (2013), e com esse nome não encontramos nenhum registro. Então partimos para uma segunda tentativa que foi buscar imagens do instrumento em catálogos de outros museus e, dessa forma, encontramos no Museu da Música de Portugal um instrumento similar com o nome: 'Yueh ch'in'. Voltamos, então, ao catálogo do MIMO e encontramos uma lista de instrumentos com nomes variados: 'Moon lute', 'Dan doan', 'Yueqin', 'Yueh Q'in'. Esse catálogo indicava como nomenclatura padrão, 'Yueqin', forma referenciada no *The New Grove Dictionary of Musical Instruments* e, portanto, resolvemos adotar esse termo padrão no nosso catálogo.

Além da nomenclatura padrão registramos nomes adicionais para cada instrumento, encontrados em diferentes fontes sempre referenciadas. De forma geral, as fontes utilizadas foram os inventários anteriores do Museu Instrumental Delgado de Carvalho, dicionários

¹¹⁵ É importante frisar que apesar de existirem todos esses inventários citados, inicialmente só tivemos acesso aos inventários de 1905 e o de 1974. Mesmo o inventário de 2008 ainda não estava disponível quando iniciamos o trabalho. Quando pudemos ter acesso aos outros inventários percebemos que, à exceção do primeiro, os outros acrescentaram pouco à nossa pesquisa em busca da nomenclatura, pois, como afirmou Almeida (1994-95), se repetiam em suas definições.

especializados e catálogos de museus com acervos de instrumentos musicais semelhantes. O registro dos nomes adicionais é importante porque os instrumentos musicais podem ser chamados por nomes diferentes dependendo do contexto e da época onde estão inseridos. A presença dos nomes adicionais possibilita que o pesquisador que não conhece a nomenclatura padrão também possa encontrar os itens do acervo que estão disponíveis no catálogo online.

Outra questão importante quando tratamos de nomenclatura é a tradução do termo padrão, em geral do inglês para português. Como mencionado anteriormente o *The New Grove Dictionary of Musical Instruments* é recomendado como fonte, porém esse dicionário é escrito em língua inglesa o que gera interpretações distintas quando se trabalha com outras línguas e outras culturas¹¹⁶. Para exemplificar essa questão, utilizaremos os exemplos dos instrumentos mostrados na FIGURA 9, denominados “Carcabas”.

FIGURA 9 – “Carcabas” do Museu Delgado de Carvalho



Na busca pelo termo padrão encontramos no sítio do MIMO o termo em língua inglesa “*Shkash*”. Porém, continuando a pesquisa por meio de conversas com um especialista da área¹¹⁷, percebemos que este não era um termo utilizado na língua portuguesa e por isso optamos por manter o termo “Carcabas”, utilizados nos inventários anteriores do Museu Instrumental Delgado de Carvalho. Dessa forma, o nome padrão “*Shkash*” está indicado no campo “Nomes adicionais”, e, além disso, no campo “Notas” (que apresentaremos mais a frente) informamos que esse é o termo padrão indicado pelo MIMO e pelo CIMCIM.

Para definirmos a classificação dos instrumentos no catálogo do MVIM, utilizamos as “Categorias”, estabelecidas pelo sistema Hornbostel & Sachs ((1914) 1961), pois é o mais utilizado em museus e bibliotecas de música.

Cada tipo de instrumento recebe um código decimal e uma descrição sucinta. O “violino”, por exemplo, é classificado como: 321.322 e, sua descrição em língua portuguesa é “Alaúdes em forma de caixa achatada com braço e cravelhame”.

¹¹⁶ Existe uma versão concisa do Dicionário Grove em língua portuguesa.

¹¹⁷ Na pesquisa sobre os idiofones e membranofones contamos com o auxílio do professor da Escola de Música da UFRJ Pedro Sá.

Para chegar a “classificação” de cada instrumento tivemos que estabelecer procedimentos. O primeiro passo foi a busca de informações no MIMO no qual chegávamos à classificação do instrumento em inglês, pois não havia, até então, um sistema em língua portuguesa. No caso do violino verificamos que sua classificação é 321.322-7 e sua descrição em inglês é “*Necked box lutes or necked guitars sounded by bowing*”.

Quando já havíamos concluído o processo de classificação tivemos acesso à publicação “Imagem es Musica” (BETHENCOURT *et al.*, 2012)¹¹⁸, que inclui uma versão da classificação Hornbostel & Sachs em português e com isso passamos a utilizar a classificação em português. No entanto, apesar dessa publicação significar um grande avanço no desenvolvimento de ferramentas de classificação dos instrumentos musicais em língua portuguesa, ainda permanecem alguns problemas: ela ainda não inclui um sétimo dígito, como no exemplo mostrada acima do violino; a descrição de cada instrumento é muito sucinta, dificultando a classificação e; os termos muitas vezes não estão muito adequados à língua portuguesa utilizada no Brasil.

Assim como no trabalho de definição da nomenclatura padrão, percebemos como é urgente a criação de ferramentas terminológicas e de classificação em língua portuguesa que possam auxiliar museus com acervos de instrumentos musicais.

Um trabalho muito importante para o catálogo é descrever cada instrumento de uma forma padronizada, apresentando suas características sempre do geral para o particular, de forma simples e objetiva, elencando traços da constituição física, acessórios, elementos decorativos, elementos técnicos, afinação. Sobre esse último quesito deve-se informar: no caso dos cordofones, a afinação das cordas; nos aerofones, a afinação e se o instrumento é transpositor ou não; nos membranofones e idiofones, se produzem som com alturas definidas ou não. Por fim, devem-se informar os eventuais danos físicos dos instrumentos musicais, tais como a falta de acessórios e de partes constitutivas. A descrição das “carcabas” pode ser vista no QUADRO 6.

QUADRO 6 – Descrição do instrumento “carcaba”

<p>Instrumento composto de dois elementos metálicos. Cada um destes elementos possui extremidades arredondadas e planas, unidas em um segmento central mais fino também feito de metal. Neste segmento central e nas extremidades das partes arredondadas existem pequenos</p>
--

¹¹⁸ Esta publicação foi realizada por pesquisadores do Departamento de Musicologia e da História da Arte Contemporânea da Universidad Complutense de Madrid em parceria com a Universidade Nova de Lisboa.

furos por onde passam finas cordas. Os dois elementos metálicos são unidos por uma das cordas localizada em uma das extremidades. O som produzido pelo instrumento não tem altura determinada. Um dos elementos de cada par está danificado em uma das extremidades arredondadas (faltam partes de metal).

Fonte: as autoras.

Um diferencial do catálogo do MVIM foi a criação de um campo que reúne curiosidades, informações históricas e usos atuais do instrumento, batizado de “para saber mais”. Esse campo não tem uma estrutura fixa, pois deve ser construído a partir das especificidades de cada instrumento. O objetivo principal é oferecer informações que demonstrem que cada instrumento musical faz parte de um contexto cultural específico, e que, portanto, tem seus usos determinados pelos grupos sociais em que estão inseridos. Estas informações devem ser dadas por meio de linguagem acessível e atraente aos leitores, de maneira que possam atingir um grande público. Um exemplo de como foi preparado o campo “para saber mais” com informações adicionais do bandolim pode ser visto no QUADRO 7.

QUADRO 7 – Exemplo do campo “Para saber mais” com curiosidades do bandolim

O bandolim deriva da mandola italiana e da mandore francesa, instrumentos equivalentes, que eram maiores com seis a oito ordens de cordas, ou seja, seis a oito pares de cordas, apesar de também existirem, na Itália e na Alemanha, mandolas menores, com quatro ou cinco pares de cordas. Na Itália, a mandola passou a ser chamada de mandolino e o instrumento tornou-se bastante popular entre os séculos XVII e XVIII. No século XIX, a família de luthiers Monzino de Milão fez várias modificações estruturais no mandolino, reforçando o corpo do instrumento e tornando-o maior. No século XIX a família Vinaccia deu continuidade às mudanças no instrumento realizadas pelos Monzinos, e construiu o bandolim moderno denominado bandolim napolitano, com quatro pares de cordas, corpo profundo, tampo harmônico em ângulo e decorado, afinação como a do violino sol²-ré³-lá³-mi⁴, e cordas tangidas por um plectro. No século XX, a família dos modernos bandolins de fundo plano continuou bastante popular, sendo utilizado no jazz, no bluegrass, na música popular irlandesa. No Brasil, o instrumento integra os conjuntos que tocam Choro, gênero musical que teve início no século XIX sendo ainda bastante popular.

Fonte: as autoras.

A “autoria ou fabricante”, o “local de fabricação” e a “data de fabricação” foram determinados por meio da pesquisa nos inventários anteriores do Museu Instrumental Delgado de Carvalho e da observação e investigação das marcas e inscrições presentes na maioria dos

instrumentos musicais. Como nem todos os instrumentos musicais têm estas informações registradas, não foi possível determiná-las para todos os itens do acervo.

Os instrumentos foram dimensionados por meio de padrões definidos internacionalmente. A unidade de peso é gramas (g) e a unidade de medida é centímetro (cm). Deve-se informar, quando possível, o comprimento máximo ou total (C) excluindo partes móveis localizadas no topo (palhetas, bocais, tudéis) e/ou na base (botões, espigões de violoncelo)¹¹⁹; a largura máxima (L)¹²⁰; a altura máxima ou total (A) excluindo partes móveis, pés (soco das harpas); e o diâmetro máximo externo (D), no caso de instrumentos cujo corpo é circular ou semi-circular.

Uma questão muito importante é a investigação das “Marcas e inscrições” presentes na estrutura física do instrumento que podem estar em etiquetas, placas metálicas, incisão, embutidos, monograma, data. As marcas e inscrições podem revelar detalhes acerca da origem dos instrumentos, tais como autor ou fabricante, local e data de fabricação. É fundamental observar se essas marcas e inscrições são anteriores ou posteriores à construção do instrumento, informando no campo com os termos “inscrição de origem” ou “inscrição posterior” seguidos das transcrições dos escritos identificados nos instrumentos.

Todos os instrumentos do Museu Instrumental Delgado de Carvalho estão sendo restaurados e esse processo está sendo registrado no catálogo, indicando as intervenções feitas, a data, e o profissional responsável pelo trabalho e, quando possível, apresentando desenhos técnicos feitos por especialistas visando tanto à preservação e recuperação do instrumento como o seu estudo científico.

Por fim, também disponibilizamos no catálogo online imagens, vídeos e áudios dos instrumentos musicais, que têm a função importantíssima de mostrar detalhes da estrutura física e dos contextos musicais nos quais os instrumentos estão inseridos.

As imagens foram feitas especialmente para o MVIM, e seguiram padrões de digitalização de imagens definidos por BAR & PFEFFERKORN: estão disponíveis no formato jpeg, guardadas em um arquivo com o número de registro do item seguido de um número seqüencial. As imagens têm os seguintes metadados: id do documento¹²¹, data de

¹¹⁹ Quando o instrumento é trapezoidal como os saltérios tomam-se duas medidas: C1=lado maior e C2=lado menor paralelo.

¹²⁰ Quando o instrumento é trapezoidal tomam-se duas medidas: L1=lado maior e L2=lado menor paralelo.

¹²¹ Adotamos para cada arquivo de imagem uma nomenclatura que contempla os seguintes dados: (1) nome do museu virtual de instrumentos musicais, (2) nome do acervo, (3) categoria, (4) número

criação, mês e ano, responsável pela criação, formato do arquivo, dimensão em pixel, tamanho do arquivo, sistema de iluminação, profundidade de cor/ resolução tonal; resolução linear (dpi ou ppi); o perfil de cor deve seguir o padrão ICC: Adobe RGB; software de captura; software de processamento.

Os registros audiovisuais, nessa primeira versão do catálogo, foram cuidadosamente selecionados do sítio *Youtube* e apresentam performances com instrumentos semelhantes aos do acervo. Está previsto para uma próxima versão do MVIM a produção de vídeos e áudios com os próprios instrumentos do acervo.

6 CONSIDERAÇÕES FINAIS

Discutimos, nesse artigo, questões relacionadas à organização e representação da informação no Museu Virtual de Instrumentos Musicais - MVIM.

Procuramos mostrar a experiência de adaptação e organização do acervo do Museu Instrumental Delgado de Carvalho no catálogo virtual do MVIM, que trouxe a tona questões relacionadas à terminologia, classificação, descrição e padronização das informações sobre instrumentos musicais.

Percebemos a importância e a necessidade de estabelecer um diálogo interdisciplinar entre a Musicologia e a Organização do Conhecimento procurando avançar mais nas questões relacionadas a padronização, descrição de acervos, vocabulários controlados, tesouros e ontologias.

Algumas iniciativas internacionais são e foram fundamentais para o nosso trabalho. Na Espanha e Portugal o projeto *Imagem es Musica* (BETHENCOURT et al, 2012) fez uma tradução da classificação Hornbostel & Sachs. Em Portugal, foram publicadas *Normas de Inventário dos Instrumentos Musicais* (TRINDADE, 2011). Na Web está disponível o projeto MIMO - *Musical Instrument Museums Online* (2013) com um vocabulário controlado, mas que ainda contempla a língua portuguesa.

Diante da necessidade de se estabelecer um vocabulário controlado em língua portuguesa na área de instrumentos musicais, iniciamos uma parceria com o Centro de Estudos de Sociologia e Estética Musical da Universidade Nova de Lisboa com o propósito de desenvolver o projeto "Designação e taxonomia dos instrumentos musicais da tradição erudita ocidental".

É importante enfatizar que o catálogo virtual do MVIM com o acervo do Museu Instrumental Delgado de Carvalho, indisponível desde 2008, possibilitará sua ampla divulgação e o início de um processo de reestruturação do Museu da Escola de Música.

Uma preocupação sempre presente no nosso trabalho foi criar um museu virtual que estivesse próximo de um público com interesses distintos. Um museu virtual que divulgue acervos de instrumentos musicais variados e que se estabeleça como um espaço aberto, dinâmico, que atraia estudantes, professores, pesquisadores, amantes da música, na qual a pesquisa e a reflexão sobre aspectos da linguagem musical, da organologia, da organização da informação e outros temas possam ser apresentados e discutidos.

REFERÊNCIAS

- BAR, Frank P.; PFEFFERKORN, Franziska. **The MIMO digitisation standard: definition of scanning properties and recommendations for photographing musical instruments.** Final version 3. Nürnberg: MIMO – Musical Instrument Museums Online, 2011.
- BESSARABOFF, Nicholas. *Ancient European Musical Instruments.* Boston: Museum of Fine Arts & Harvard University Press, 1941.
- BETHENCOURT *et al.* **Imagem es musica: Proyecto Iconografía Musical UCM y CESEM/NIM,** Universidade Nova de Lisboa. Madrid: AEDOM, 2012. 1 DVD.
- BRANDÃO, Dolores Castorino. **Representação documentária de instrumentos musicais: contribuição para a organização do Museu Instrumental Delgado de Carvalho da Escola de Música da Universidade Federal do Rio de Janeiro,** 2013 (Monografia de Especialização). Arquivo Nacional - UFRJ.
- CAMPOS, Maria Luiza Almeida. **A organização de unidades do conhecimento em hiperdocumentos: o modelo conceitual como espaço comunicacional para a realização da autoria.** 2001a. Tese (Doutorado em Ciência da Informação). Programa de Pós-graduação em Ciência da Informação, CNPq/IBICT - UFRJ/ECO.
- CARDIM, Fernão. **Tratados da terra e gente do Brasil.** 3. ed. São Paulo: Comp. Ed. Nacional; Brasília: INL, 1978.
- CARDOSO, André. A Escola de Música e suas coleções especiais. In: Oliveira, Antônio José Barbosa de (Org.). **Universidade e lugares de memória.** Rio de Janeiro: UFRJ/FCC/SIBI, 2008. p. 203-220.
- CARVALHO, Delgado de. *O Museu Instrumental do Instituto Nacional de Música do Rio de Janeiro.* Rio de Janeiro: Imprensa Nacional, 1905.
- DAHLBERG, Ingetraut. Classification theory, yesterday and today. **International Classification**, v. 3, n. 2, p. 85-90, 1976.
- DAHLBERG, Ingetraut. **Definitionen aus dem Begriffsfeld Wissensorganisation: vorgelegt von Dr. Ingetraut Dahlberg, Bad König.** 2006. Disponível em: <http://www.db.dk/bh/Lifeboat_KO/CONCEPTS/knowledge_organization_Dahlberg.htm>. Acesso em: jan. 2009.

DODEBEI, Vera L. D. **Tesouro**: linguagem de representação da memória documentária. Niterói; Rio de Janeiro: Intertexto; Interciência, 2002.

DOURNON, Genevieve. Instrumentariums et classifications. **Revue de Musicologie**, T. 79e, n. 2e, p. 297-307, 1993.

DUCKLES, Vincent; BALCHIN, Robert. **Musicology**. Grove Music Online. Oxford Music Online. Disponível em: <<http://www.oxfordmusiconline.com/subscriber/article/grove/music/46710pg2>>. Acesso em: fev. 2009

GROVE. **Music Online**. Oxford Music Online. Disponível em: <http://www.oxfordmusiconline.com/public/book/omo_gmo>. Acesso em: nov. 2013.

HOLOMAN, D. Kern; PALISCA, Claude V. (Org.). **Musicology in the 1980s**: methods, goals, opportunities. New York: Da Capo Press, 1982.

HORNBOSTEL, Erich M. Von; SACHS, Curt Sachs. Systematik der Musikinstrumente. Ein Versuch, vol. xlvi, 1914, p.553-590. Classification of Musical Instruments: Anthony Baines and Klaus P. Wachsmann (trad.). **The Galpin Society Journal**, v. 14, p. 3-29, Mar. 1961.

KARTOMI, Margaret. The Classification of Musical Instruments: Changing Trends in Research from the Late Nineteenth Century, with Special Reference to the 1990s. **Ethnomusicology**, v. 45, n. 2 (Spring - Summer, 2001), p. 283-314.

LANZELOTTE, Rosana; ULHOA, Martha Tupinambá de; BALLESTÉ, Adriana Olinto. Sistemas de informações musicais: disponibilidade de acervos musicais via Web. **Opus - Revista da Associação Nacional de Pesquisa e Pós-Graduação em Música**, Campinas, n. 10, 2004.

MENEZES, Antonio Cardoso de. [Crítica]. **Gazeta Musical**, v. 2, p. 37, 1892.

MIMO - Musical Instrument Museums Online. Disponível em: <<http://www.mimo-db.eu/MIMO/Infodoc/>>. Acesso em: novembro de 2013.

MIRANDA, Marcos L. C. de. **Organização e representação do conhecimento**: fundamentos teórico-metodológicos na busca e recuperação da informação em ambientes virtuais. (Doutorado em Ciência da Informação). Convênio MCT/IBICT-UFRJ/ECO, Rio de Janeiro, 2005.

SACHS, Curt. **The history of music instruments**. (New York: Norton, 1940). London: J.M. Dent & Sons, 1968.

SCHAEFFNER, André. **Origine des instruments de musique**. Réimpression augmentée. Paris: Payot, 1936.

TRINDADE, Maria Helena. **Normas de inventário**: instrumentos musicais. Instituto dos Museus e da Conservação, 2011.

SEMIÓTICA DA CULTURA E ABORDAGENS SOCIOCULTURAIS: POSSÍVEIS DIÁLOGOS

SEMIOTICS CULTURE AND SOCIO-CULTURAL APPROACHES: POSSIBLE DIALOGUE

Mona Cleide Quirino da Silva Farias
Carlos Cândido de Almeida

Resumo: A Semiótica da Cultura é um campo de estudo cujas investigações estão para análise da linguagem, mas precisamente das linguagens da cultura. A concepção de linguagem e cultura são elementos que impulsionam as discussões na Semiótica da Cultura. Buscamos a reflexão em torno da diversidade cultural e da diversidade linguística existente em diversos contextos da cultura. A questão que norteou a análise foi saber quais as contribuições a Semiótica da Cultura pode oferecer ao campo da Organização do Conhecimento. Objetivamos investigar conceitos das abordagens socioculturais da Organização do Conhecimento por uma perspectiva teórico-conceitual da Semiótica da Cultura. Assim, buscamos traçar uma análise sobre relações teórico-conceituais entre a Semiótica da Cultura e as abordagens socioculturais com destaque aos autores: Begthol, García Gutiérrez e Hudon. A pesquisa possui caráter exploratório e abordagem qualitativa. O *corpus* reuniu materiais do campo da Organização do Conhecimento e da Semiótica da Cultura. O método de análise foi a *Análise de Conteúdo*. Realizamos as análises sobre os conceitos que caracterizam as abordagens socioculturais e as concepções da Semiótica da Cultura. Após a análise apresentamos o diálogo entre a Semiótica da Cultura e as abordagens socioculturais. Consideramos que os objetivos foram alcançados. Portanto, entendemos que a Semiótica da Cultura pode contribuir com a Organização do Conhecimento com base em uma análise da cultura e dos contextos das comunidades, tendo em vista diálogos interdisciplinar.

Palavras-chave: Semiótica da Cultura. Diversidade cultural. Diversidade linguística. Abordagens Socioculturais. Organização do Conhecimento.

Abstract: The Semiotics of Culture is a field of study whose investigations are to analyze the language, but precisely the languages of culture. The conception of language and culture are driving the discussions in the Semiotics of Culture. Nicer reflection on cultural diversity and existing in various contexts of culture linguistic diversity. The question that guided the analysis was to know which contributions Semiotics of Culture can offer to the field of Knowledge Organization. We aimed to investigate concepts of sociocultural approaches to Knowledge Organization for a theoretical-conceptual perspective of Semiotics of Culture. Thus, we seek to draw an analysis of theoretical and conceptual relationships between the Semiotics of Culture and sociocultural approaches highlighting the authors: Begthol, García Gutiérrez and Hudon. The research has exploratory and qualitative approach. The corpus gathered materials from the field of Knowledge Organization and the Semiotics of Culture. The analysis method was content analysis. We performed analyzes of the concepts that characterize the socio-cultural approaches and conceptions of Semiotics of Culture. After analyzing the present dialogue between the Semiotics of Culture and sociocultural approaches. We believe that the objectives were achieved. Therefore, we believe that the Semiotics of Culture can contribute to the Knowledge Organization based on an analysis of culture and contexts of the communities with a view interdisciplinary dialogues.

Keywords: Semiotics of Culture. Cultural diversity. Linguistic diversity. Sociocultural approaches. Knowledge Organization.

1 INTRODUÇÃO

Reconhecemos as abordagens socioculturais¹²² como um dos núcleos de estudo da Organização do Conhecimento¹²³ que contempla estudos, cujas investigações direcionam para um olhar sobre a cultura e o contexto. O objetivo da pesquisa foi investigar conceitos das abordagens socioculturais da Organização do Conhecimento por uma perspectiva teórico-conceitual da Semiótica da Cultura.

Denominamos abordagens socioculturais da Organização do Conhecimento uma linha de estudos que está orientada por discussões sobre os aspectos éticos, culturais, representacionais e linguísticos. Entre os teóricos podemos citar: García Gutiérrez (1998, 2002), Beghtol (2002), Guimarães e Fernández-Molina (2003), Hudon (1997), Hjørland (2007, 2013), Campbell (2000, 2010), entre outros.

Os estudos das abordagens socioculturais analisados na pesquisa tratam especificamente das seguintes temáticas: da “Garantia cultural e hospitalidade cultural” de Clare Beghtol; “Ética transcultural da mediação” de Antônio Garcia Gutiérrez e “Tesauros multilíngues” de Michèle Hudon. No contexto dessas discussões, entendemos a cultura como um elemento de reflexão no âmbito da Organização do Conhecimento, a qual se encontra em destaque pelas investigações propostas nas abordagens socioculturais desse campo.

Consideramos que a cultura pode se tornar um elemento de significativa reflexão para a Organização do Conhecimento, se analisada com o objetivo de conhecer os contextos culturais.

A Semiótica da Cultura, também conhecida como Semiótica Russa, desenvolveu suas discussões sobre os sistemas de signos¹²⁴, a cultura¹²⁵, o texto cultural¹²⁶, códigos culturais¹²⁷

¹²² As abordagens socioculturais são investigações, cujos estudos revelam atenção para aspectos éticos e socioculturais no contexto da Organização do Conhecimento e desse modo contribuem para análise das questões éticas, políticas, linguísticas, sociais e culturais que se direcionam para a configuração de sistemas de organização e representação do conhecimento.

¹²³ Organização do Conhecimento atua como um campo propício de reflexão para os contextos culturais, tendo como ponto de partida as abordagens socioculturais. A Organização do Conhecimento pode ser compreendida enquanto campo de mediação e, nesse sentido, podemos dizer que a linguagem assim como a cultura são conceitos que transitam nesse espaço.

¹²⁴ A concepção de sistemas de signos busca dar ênfase aos sistemas sógnicos constituídos por signos que transcendem o caráter verbal como, por exemplo, os signos não-verbais.

¹²⁵ A cultura é tratada como um conjunto de informações não-hereditárias, isto é, de conteúdos armazenados e comunicados por grupos distintos. A cultura, para a Semiótica eslava é fruto da semiose, isto é, da própria natureza (MACHADO, 2010). De modo geral, a cultura é um sistema semiótico caracterizado por sua capacidade dinâmica e transformadora sobre os textos criados e desenvolvidos por uma memória coletiva (MACHADO, 2003, p. 157).

¹²⁶ Consiste em mecanismos semióticos e atuam na geração de sentidos. São produções da cultura e interagem com o contexto.

e linguagens¹²⁷, em meados da década de 1960. Entendemos que tais discussões podem se somar às discussões sobre linguagem e cultura que potencializem e contribuam com os estudos contemplados pelas abordagens socioculturais, a saber, as questões que envolvem a linguagem, os aspectos éticos socioculturais da Organização do Conhecimento.

O contexto no qual a Semiótica da Cultura emerge foi a Escola de Tártu-Moscou (ETM), na Universidade de Tártu, Estônia. Dos pioneiros responsáveis pelas discussões teóricas nesse campo de estudos da linguagem, podemos citar Yuri, M. Lotman (1999, 1990, 2003) com sua análise sobre semiosfera. A semiosfera é o espaço de desenvolvimento dos sistemas de signos da cultura, isto é, um espaço semiótico sujeito a processos de modelização¹²⁹, e transcodificação¹³⁰, um contínuo semiótico.

Na Semiótica da Cultura, também é destacado a concepção de linguagens. As linguagens são processos comunicativos de múltiplas expressões. Elas revelam o contexto híbrido (CANCLINI, 2008) das culturas e, por conseguinte, dos sujeitos que fazem parte destas atuando como agentes produtores de textos manifestados por múltiplas linguagens. Desse modo, compreendemos que “[...] todas as práticas humanas são tipos de linguagens, já que elas têm a função de demarcar, significar e comunicar” (CINTRA, et al., 2002, p.26). A linguagem¹³¹ é um modo de expressão podendo ser expressa por diversos signos que não seja somente o signo linguístico (MACHADO, 2003, p. 163) e no contexto da Semiótica da Cultura é um elemento a que devemos destaque. Em torno da discussão de linguagens, nossas reflexões também vão ao encontro da perspectiva de uma semiótica social, na qual considera a

¹²⁷ São estruturas complexas que reconhecem, armazenam e processam informações. Também constituem um vocabulário mínimo e dinâmico da cultura. São sistemas modelizantes, pois tratam de uma forma de regulação necessária para a organização e desenvolvimento da informação (MACHADO, 2003, p. 156).

¹²⁸ Nesse sentido, a linguagem é um dispositivo semiótico que pode ser expresso por meio da arte, de uma imagem, de um som, etc.

¹²⁹ Conceito atribuído pelos semioticistas no campo da Semiótica da Cultura, cujo propósito foi buscar a distinção entre a língua natural enquanto sistema primário dos demais sistemas de linguagens culturais, que na Semiótica da Cultura, se considera, por exemplo, a música, a dança, a poesia etc., tendo em vista que ambos se originam de um sistema modelizante primário, com base em uma estrutura, a língua.

¹³⁰ O termo transcodificação corresponde a “operação” (ou conjunto de operações) pela qual um elemento ou um conjunto signifiante é transposto de um código para outro, de uma linguagem para outra” (GREIMAS; COURTÉS, 2008, p. 509). Para a Semiótica da Cultura, por exemplo, os códigos culturais são denominados processos de transcodificação, e por isso não se trata de uma mera decodificação (MACHADO, 2003).

¹³¹ Segundo Semprini (1999), a linguagem é um instrumento cultural potencializador do conhecimento humano e por isso interfere na visão de mundo dos sujeitos.

própria linguagem enquanto “semiótica social” (HALLIDAY, 1982, p. 9) revelada nos contextos¹³² da sociedade.

Além das concepções de cultura, linguagem, modelização, outro termo explorado pela Semiótica da Cultura é a tradução, esta consiste em um exercício interpretativo sobre os aspectos e elementos culturais e não condiz em uma mera reprodução. A tradução¹³³, por sua vez, é tratada aqui enquanto uma atividade semiótica que envolve necessariamente a interpretação e compreensão no objetivo de representar uma determinada realidade cultural.

Para a Organização do Conhecimento, tradução é entendida como uma passagem de uma língua para outra, no caso, da língua dos documentos para a língua do sistema. No contexto das Linguagens Documentais, por exemplo, a tradução “[...] é feita em unidades informacionais ou conjunto de unidades aptas a integrar sistemas documentários” (CINTRA *et al.*, 2002, p.35).

Mediante as concepções de cultura, linguagem, texto, contexto, tradução e modelização desenvolvidas pela Semiótica da Cultura, consideramos sua pertinência, e entendemos que trazem uma perspectiva ampla de compreensão da diversidade linguística e cultural de determinados domínios do conhecimento¹³⁴, que nos leva à necessidade de destacar que há uma tendência em se discutir sobre as abordagens socioculturais na Organização do Conhecimento, principalmente pela constatação de determinados estudos desenvolvidos, o que nos impulsiona a compartilhar com demais pesquisas que possam vir a somar conteúdos aos estudos socioculturais da Organização do Conhecimento.

Buscamos aproximações de ordem teórica e conceitual com os estudos da Semiótica da Cultura. Portanto, afirmamos de antemão, que o estudo da linguagem, visto por um viés semiótico, pode trazer contributos significativos para Ciência da Informação e de maneira mais específica ao núcleo de discussões socioculturais da Organização do Conhecimento, pois consideramos que este campo, enquanto um espaço que desenvolve processos de mediação

¹³² Representam o momento histórico, social e cultural vivenciado pelas comunidades e culturas. Também, mencionamos sobre a concepção de texto cultural, ou seja, as produções culturais manifestadas por linguagens, enquanto criações dialógicas no contexto das culturas.

¹³³ A tradução sobre a qual refletimos nesta pesquisa é uma tradução de caráter intersemiótica ou de transmutação, isto é, na interpretação dos signos verbais via sistemas de signos não-verbais (JAKOBSON, 1995, p. 65). A tradução é um mecanismo que possibilita ultrapassar uma interpretação que esteja baseada somente em signos verbais. Por exemplo, a tradução de elementos simbólicos no contexto cultural pode apresentar uma variedade de elementos passíveis para análises interpretativas.

¹³⁴ Campos de atividades que compartilham de conhecimentos que se aproximam, tendo em vista interesses comuns como, por exemplo, a Semiótica da Cultura e os campos de investigações da Antropologia, Linguística estrutural, entre outras.

precisa cada vez mais estender seus estudos e reflexões, tendo em vista dialogar com diversas perspectivas culturais e da linguagem que lhes favoreçam uma análise interdisciplinar sobre as questões socioculturais, da cultura e do contexto das comunidades.

A partir das reflexões desencadeadas pela Semiótica da Cultura, supomos a hipótese de um diálogo entre a Semiótica da Cultura e as abordagens socioculturais, no intuito de destacar as relações de ordem teórica e conceitual que mais se aproximem das análises sobre linguagem e cultura e que possam proporcionar uma reflexão por meio das abordagens socioculturais¹³⁵.

Destacamos que há a necessidade em discorrer sobre a concepção de linguagens, e de sua função mediadora mediante as abordagens socioculturais na Organização do Conhecimento, considerando uma perspectiva ampla de linguagem, pois entendemos que é preciso atentar para o pluralismo linguístico existente nos diversos contextos culturais dos usuários que recorrem a um sistema de representação da informação em busca de um conteúdo que corresponda ao seu contexto.

Procuramos investigar os elementos conceituais comuns entre os dois campos de discussão que é a Semiótica da Cultura com suas revelações quanto à cultura, linguagem, texto, contexto, tradução e sistemas modelizantes, e a Organização do Conhecimento com as abordagens socioculturais e suas investigações focadas nos aspectos culturais.

A pesquisa pretende contribuir significativamente para o estudo das abordagens socioculturais desenvolvidas no contexto da Organização do Conhecimento, na perspectiva de revelar possibilidades de diálogos interdisciplinares com este campo de estudo. Ademais, consideramos a relevância desta pesquisa pelo fato da mesma trazer elementos de comparação com demais pesquisas na área da Ciência da Informação, principalmente no campo da Organização do Conhecimento. Nesse sentido, a pesquisa busca compartilhar com as abordagens socioculturais especificamente em análises das questões socioculturais.

Sustentamos que ao aproximarmos as concepções da Semiótica da Cultura em relação aos conceitos e perspectivas das abordagens socioculturais da Organização do Conhecimento,

¹³⁵ É relevante mencionar que dos estudos desenvolvidos na Organização do Conhecimento, mais precisamente aqueles relacionados às abordagens socioculturais, Barát (2008), Beghtol (1986, 1995, 2001, 2002 e 2005), Bravo (2006), Campbell (2002, 2010), Dal' Evedove (2014), García Gutiérrez (1998, 2002a, 2002b, 2002c, 2004, 2006 e 2008), Guimarães (2005); Guimarães e Milani (2010, 2011), Guimarães e Pinho (2008), Hjørland (2003, 2007, 2010, 2013), Hudon (1997, 1999, 2003), López-Huertas (2006, 2008, 2010), Milani (2010), Moura (2010), Pinho (2006, 2010) entre outros, embora tenham discutido sob o ponto de vista sociocultural, não tiveram como objetivo apresentar a relação da Semiótica da Cultura com as abordagens socioculturais da Organização do Conhecimento.

podemos traçar posteriores análises da cultura, da linguagem cujo objetivo seja a abertura para análises com caráter interdisciplinar no campo da Organização do Conhecimento e de modo mais abrangente para a área da Ciência da Informação¹³⁶.

A pesquisa se justifica pertinente por sua realização estar direcionada a uma análise da concepção da linguagem e cultura justificando a importância em atentar para a análise dos contextos culturais das comunidades. A partir disso, buscamos identificar e discutir sobre as possibilidades de intersecção entre a Semiótica da Cultura e o núcleo de estudo que contempla os aspectos socioculturais na Organização do Conhecimento. Vale destacar que o desenvolvimento desta pesquisa partiu da necessidade de verificar diálogos entre a Semiótica da Cultura e a Organização do Conhecimento. O interesse foi favorecer uma análise crítica que aponte o grau de interação e o diálogo entre os dois campos de estudo.

2 ORGANIZAÇÃO DO CONHECIMENTO: ABORDAGENS SOCIOCULTURAIS

A Organização do Conhecimento¹³⁷ é um campo híbrido, por agregar uma diversidade de temas e estudos com abordagens distintas e que demonstram potencial para diálogos interdisciplinares. Este campo exige dos profissionais envolvidos atenção no desenvolvimento de processos de organização e representação, cujo objetivo é articular processos críticos e reflexivos sobre todo o conhecimento criado, em destaque do caráter mediador desse campo.

Entendemos que a Organização do Conhecimento enquanto campo de estudo interdisciplinar ganha concepções diversas. Muitos a consideram como ciência, cujo objetivo é a estruturação e sistematização de conceitos (DAHLBERG, 2006); outros a consideram como processo de organização da informação em registros bibliográficos, no âmbito da Biblioteconomia e da Ciência da Informação (HJORLAND, 2003); vista ainda como disciplina vinculada à Ciência da Informação, tendo em vista os fundamentos teóricos do tratamento e a recuperação da informação (ESTEBAN NAVARRO, 1996). Considerada ainda como ramo do conhecimento, tendo em vista a construção de ferramentas que auxiliem no armazenamento de entidades documentárias (SMIRAGLIA, 2002); numa perspectiva do

¹³⁶ A Semiótica foi objeto de outros estudos na área da Ciência da Informação, como podemos verificar em Almeida (2010), Mai (1997a, 1997b, 2001 e 2004) Monteiro (2006a e 2006b), Moura (2005, 2006b, 2007 e 2011), Hjørland (2007), Sousa e Almeida (2012), Thellefsen (2002, 2003, 2004a, 2004b e 2009), Thellefsen e Thellefsen (2004) entre outros, mas nesses estudos não conseguimos evidenciar o enfoque especificamente da Semiótica da Cultura. Estas fontes só reforçaram o desenvolvimento desta pesquisa, cogitando nos contributos que a mesma poderá dispor tanto para a Organização do Conhecimento como à área da Ciência da Informação.

¹³⁷ Enfatizamos que quando a expressão Organização do Conhecimento é destacada em maiúscula estamos tratando do campo e quando em minúscula direcionamos nossas reflexões aos processos de organização e representação do conhecimento.

ensino em Biblioteconomia, a Organização do Conhecimento a qual abarca a área do Tratamento Temático da Informação (GUIMARÃES; PANDO, 2006), entre tantas outras abordagens.

Guimarães e Dobedei (2012, p. 13-14) comentam que a Organização do Conhecimento “encontra-se diretamente ligada a International Society for Knowledge Organization – ISKO¹³⁸, criada em 1989, na Alemanha tendo à frente Ingetraut Dahlberg”.

Fujita (2008) destaca que a “ISKO é a sociedade científica da área de Organização e Representação do Conhecimento responsável pelas principais ações em torno de sua necessária consolidação científica”. Segundo Guimarães e Dobedei (2012, p. 15), existem atualmente 12 capítulos da ISKO em atividade, por exemplo, Brasil, Canadá/Estados Unidos, China, França, Alemanha/Áustria/Suíça, Índia, Itália, Maghreb (Argélia, Marrocos e Tunísia), Nórdico incluindo (Dinamarca, Suécia, Noruega, Finlândia, Islândia e Faroé), Polônia, Espanha e Reino Unido (ISKO, 2013).

Dahlberg (2006, tradução nossa) relata que a Organização do Conhecimento é uma ciência¹³⁹ pela qual é possível ordenar a estruturação e sistematização dos conceitos. No mesmo trabalho compreende que o conhecimento possui uma natureza subjetiva e individual, intransferível. Ao compreender o “Conhecimento em ação” tal como destacado por Dahlberg (1993) também consideramos “a cultura, a linguagem em ação” as quais estão imbuídas na configuração do conhecimento.

Em suma, a Organização do Conhecimento é um campo que possibilita mediações entre o conhecimento representado e a comunidade usuária, tendo em vista a atuação crítica por parte dos profissionais da informação e seu papel social em consideração à diversidade cultural presente nas comunidades. O campo da Organização do Conhecimento estuda os domínios do conhecimento com suas peculiaridades, isto é, em que se distinguem de outros, pois o conhecimento não é algo único e fechado.

O domínio do conhecimento é entendido como uma demarcação de determinado conhecimento, ancorado ou não em um ambiente profissional. Dessa maneira, domínio do conhecimento é um tipo de significado que organiza o conhecimento em relação a uma área específica sob uma determinada perspectiva. É sob essa perspectiva que se compreende a

¹³⁸ <http://www.isko.org/about.html>

¹³⁹ Dahlberg (1993, p. 2011, tradução nossa), relata que a Organização do Conhecimento é entendida como uma “ciência que estrutura e também organiza de modo sistemático unidades do conhecimento, isto é, conceitos, considerando assim seus elementos de conhecimento, características, as quais são inerentes à aplicação de conceitos e classes de conceitos ordenados a objetos/assuntos”.

informação a partir de seu contexto ou de quem a produz (PINHO, 2010, p. 50).

Compreendemos a Organização do Conhecimento como um campo de ampla discussão, onde podemos desenvolver diálogos e reflexões que poderão desencadear metodologias que colaborem com este campo tendo em vista refletir não apenas processos de organização, mas também sobre o conhecimento que emerge da dinâmica das comunidades.

Uma das dez premissas básicas do campo da Organização do Conhecimento apresentadas por Barité (2001, p. 42 tradução nossa) considera o conhecimento um produto social e uma necessidade social, ressalta que “[...], em suma, pode ser visto como o saber acumulado da humanidade, mas também fornecedor permanente que atenda às necessidades sociais e identifica “lacunas” onde novos conhecimentos são necessários para se regenerar”.

Entendemos o conhecimento¹⁴⁰ como uma operação organizada e dialógica em constante processamento, se trata de algo construído não somente pelo indivíduo, mas pelas relações com o coletivo. Sendo assim, entendemos que o caráter dialógico do conhecimento requer a compreensão do contexto no qual foi criado. O conhecimento pode ser descrito como a soma de experiências, organizadas e sistematizadas, visto que tal conhecimento não é algo acabado e por isso devemos considerar que em cada contexto, por ocasião das relações estabelecidas entre os sujeitos, este conhecimento possuirá características distintas e estará sempre se renovando, em constante desenvolvimento.

Vale destacar que se o conhecimento é entendido enquanto uma ação social é necessário mencionar que no âmbito social tal conhecimento não está desprovido de anuências. Guimarães (2001) afirma que o conhecimento é “algo sobre o qual existe certo consenso social [...]”. Podemos perceber o mesmo, como algo complexo e muitas vezes de difícil compreensão, visto que depende de um consenso no âmbito dos contextos sociais.

De modo geral, necessitamos organizar e representar os conhecimentos no intuito de mediar os demais conhecimentos para acesso à informação. Para tanto, vale mencionar sobre a representação do conhecimento. Brascher e Café (2008, p. 6) destacam que a representação do conhecimento trata de temáticas que nos encaminham ao “mundo dos conceitos” e não meramente ao dos registros de informação. A representação do conhecimento busca refletir

¹⁴⁰ O conhecimento como um produto da sociedade compõe domínios diversos, na medida em que este pode ser constituído por uma determinada comunidade. A partir disso, consideramos que o conhecimento enquanto ação social está sujeito a constantes mudanças às quais podem ocorrer pelo fato de os sujeitos estarem em contato com uma diversidade cultural e linguística.

uma visão consensual sobre a realidade a ser representada. Em outras palavras é um processo reflexivo sobre os contextos em que estão inseridos diversos domínios de conhecimento.

Segundo Campos (2004, p. 24), a representação do conhecimento é em si um mecanismo de raciocínio sobre o mundo. Um modo de reflexão sobre o mundo sem precisar agir diretamente sobre ele. A autora acrescenta que a representação

[...] é, fundamentalmente, um substituto para aquilo que representa. [...]. Deve haver alguma forma de correspondência específica entre o substituto e seu referente planejado no mundo. Esta correspondência é a semântica da representação. [...]. A única representação completamente precisa de um objeto é o objeto em si. Qualquer outra representação é imprecisa e, inevitavelmente, contém simplificações (CAMPOS, 2004, p. 24).

Enquanto processo, a organização do conhecimento implica a reflexão por parte dos profissionais envolvidos para a mediação do conhecimento de modo que seja possível atentar para os contextos socioculturais das comunidades.

As abordagens socioculturais¹⁴¹ representam investigações que ressaltam a necessidade de compreender a cultura e o contexto das comunidades discursivas. É necessário destacar que há estudos que enfatizam abordagens socioculturais. Um exemplo que podemos mencionar trata da tese desenvolvida por Dal' Evedove (2014,) que destaca uma discussão sobre abordagem sociocultural. Nesse trabalho a autora resgata a abordagem sociocultural no campo da Informação considerando que esta é uma alternativa para observar as atividades informacionais tendo como base as pessoas e os seus contextos.

Com base nas abordagens socioculturais da Organização do Conhecimento, mais precisamente as reflexões de Beghtol (2002), Hudon (1997, 1999) e García Gutiérrez (1998, 2002a, 2002b, 2002c, 2004 e 2008), objetivamos apresentar os seguintes temas: garantia cultural, hospitalidade cultural, multilinguismo e ética transcultural da mediação, que contemplam reflexões de ordem sociocultural na Organização do Conhecimento.

As abordagens socioculturais da Organização do Conhecimento revelam a necessidade de reflexão sobre a cultura, tendo em vista a diversidade cultural e demais aspectos que envolvem o conhecimento, assim como sua representação. Desse modo, refletem sobre a linguagem, tendo em vista a questão da postura ética e de mediação dos profissionais da informação. Em outras palavras, buscam priorizar a diversidade cultural de indivíduos e enfatiza para isso a necessidade em compreender os contextos das comunidades.

¹⁴¹ Estas investigações que desencadeiam questões de ordem cultural, ética e multilíngue são consideradas relevantes no contexto de discussão desta pesquisa. Entendemos que podem auxiliar a refletir sobre os processos de organização e representação do conhecimento sob uma vertente sociocultural.

As abordagens socioculturais da Organização do Conhecimento representam investigações que ressaltam a necessidade de atenção para a cultura e o contexto das comunidades, tendo em vista a socialização do conhecimento. A Organização do Conhecimento com o enfoque nos estudos socioculturais nos alertam para a prática do profissional¹⁴² enquanto sujeito mediador no processo de organização e representação do conhecimento. No campo da Organização do Conhecimento as abordagens socioculturais representam uma ampliação do ponto de vista teórico destacando o papel social deste campo e dos sistemas de organização e representação do conhecimento.

3 SEMIÓTICA DA CULTURA: CAMPO E OBJETO

A Semiótica da Cultura é uma teoria que busca a análise das linguagens existentes entre a natureza e a cultura (MACHADO, 2003, p. 25). A Semiótica da Cultura possui um campo transdisciplinar que dialoga com diversas disciplinas como Teoria Literária, Linguística estrutural, Semiótica, Crítica da Arte, Cibernética, Teoria da Informação e da Comunicação, Antropologia, Neurolinguística, Etnologia entre outras.

Os estudos da Semiótica da Cultura, mais precisamente a perspectiva traçada pela abordagem russa, trata de discussões sobre a linguagem na cultura¹⁴³. A linguagem é considerada como um dispositivo que se expressa por signos que não possuem relação direta com os signos linguísticos.

A Semiótica da Cultura busca entender qual papel a linguagem exerce na cultura. A ideia de texto surge com a intenção de compreender a cultura de modo mais abrangente, isto é, como uma maneira em que transitam códigos culturais e linguagens diversas. A língua natural, enquanto instituição social (SAUSSURE, 1970), é considerada como uma possibilidade de criação de linguagens, por isso compreendemos que não podemos deixar de ressaltar seu caráter comunicativo. O texto é um dispositivo que constitui uma diversidade de códigos e conseqüentemente de linguagens. Os textos caracterizam-se por mecanismos semióticos geradores de sentido (MACHADO, 2003).

¹⁴² Guimarães (2005) ressalta a “dimensão profissional da ética enquanto conjunto de valores que um segmento social caracterizado pela especificidade de um saber e de um fazer (profissão) estabelece como necessários e fundamentais ao exercício dessa profissão [...]”. Assim, compreendemos que no âmbito da Organização do Conhecimento já se evidencia um cuidado com questões que envolvem o fazer do profissional assim como sua atuação enquanto mediador.

¹⁴³ Pode ser entendida também como um processo semiótico, onde as narrativas constituídas a partir das linguagens expressam o complexo de signos. Por meio dessas narrativas podemos perceber como ocorrem os processos interpretativos e de apropriação da cultura, isto é, a recepção e mediação de seus elementos simbólicos assim como suas linguagens.

O propósito norteador da Semiótica da Cultura é compreender a comunicação como um processo semiótico enquanto que a cultura é entendida como um conjunto unificado de sistemas sógnicos constituídos por uma dinâmica de códigos e linguagens operados num contínuo semiótico. A cultura, nessa abordagem, é compreendida como um grande texto. Este é um aspecto elementar da semiótica moderna, principalmente pelo seu caráter dinâmico e dialógico.

O objeto da Semiótica da Cultura consiste nos sistemas semióticos da cultura desenvolvidos na própria, isto é, um espaço semiosférico. Contudo, entende-se a cultura como um texto, criado a partir da conceptualização de códigos culturais e linguagens. Portanto, seu desenvolvimento se dá num espaço semiótico, isto é, num espaço propício às transformações, um contínuo semiótico.

Compreendemos que o objeto da Semiótica da Cultura ganha uma dimensão maior por não tratar apenas da cultura, de suas características, instrumentos específicos, tipos de cultura. A ideia ou “cognição” como diria na semiótica moderna (NÖTH, 1996), trazida pelos estudos da Semiótica da Cultura entende a cultura como um texto e enfatiza o mesmo como um espaço semiótico composto numa semiosfera, em que são produzidos e organizados os códigos, as linguagens, assim como as semioses¹⁴⁴, isto é, os dispositivos pensantes da cultura.

Assim, seria necessário uma reflexão sobre os sistemas de signos na própria cultura, isto é, de diversos outros textos que poderíamos tratá-los como textos com dimensões sógnicas menores (não por possuírem menor sentido) no espaço da cultura, mas tratam de criações com “peculiaridades substanciais” (SCHNAIDERMAN, 1979), que no instante da análise as interpretações serão diversas, porém requerem a compreensão prévia de um repertório de signos¹⁴⁵. A ação do signo extrapola a linguagem verbal por esta constituir de uma estrutura, a língua¹⁴⁶. As linguagens não-verbais¹⁴⁷ possuem tipos de unidades diversas como, por

¹⁴⁴ Nesse sentido, a semiose é entendida como: “O processo pelo qual algo funciona como signo podendo ser chamado de semiose” (MORRIS, 1976, p. 13).

¹⁴⁵ Peirce (1990) afirma que o signo é tudo que sob certo aspecto ou modo representa algo para alguém, ou ainda, dirige-se a alguém, podendo criar na mente de uma pessoa um signo equivalente ou até mesmo um mais desenvolvido. Nesse sentido, pensamos que o signo requer processos interpretativos, assim também ocorre com o texto, mais precisamente com o texto cultural na Semiótica da Cultura.

¹⁴⁶ Vale mencionar que a língua, do ponto de vista da linguística, constitui um sistema de signos estruturados, que permitem a transmissão de mensagens por dois elementos: o “significante e significado” (SAUSSURE, 1970).

¹⁴⁷ As linguagens não-verbais são a própria representação de ideias, assim como de pensamentos, cujo objetivo é realizar a comunicação.

exemplo, o gesto, a dança, a imagem, o movimento, etc. As linguagens não-verbais não se limitam a uma unidade como a palavra, por exemplo. Vale dizer que as linguagens não-verbais como, por exemplo, o som, a imagem, e outros, também compartilham de uma estrutura que pode ser identificada pelos traços, as cores, etc.

Contudo, o texto cultural a partir de processos interpretativos, isto é, de processos modelizantes é desenvolvido no espaço da cultura, tendo como fio condutor as diversas formas de linguagem. Portanto, a ideia é identificar e compreender uma variável simbólica na cultura, campo semiótico, e esta variável deve ser destacada no intuito de poder discorrer sobre as linguagens híbridas como, por exemplo, imagem, som, indumentária, etc.

Vale ressaltar que, no âmbito de discussões da Semiótica da Cultura não se considera a estruturalidade¹⁴⁸ da língua, mas a estruturalidade das linguagens criadas e desenvolvidas na cultura, isto é, no espaço semiótico. Desse modo, a língua não serve como modelo para análise dos sistemas de signos na cultura; pelo contrário, atua apenas como um mecanismo pelo qual é possível desenvolver linguagens.

Destacamos a importância das linguagens que transcendam a linguagem verbal como, por exemplo, linguagem visual, sonora e gestual. Para tanto, não podemos deixar de considerar o contributo exercido por teóricos da linguística como Saussure (1970), assim como de Jakobson (1995), sobre sistemas de signos constituídos a partir de uma estrutura (língua).

A Semiótica da Cultura é considerada uma teoria sistêmica cujas discussões se encontram inclinadas para os sistemas de signos produzidos em sociedade, isto é, aos textos produzidos pelos indivíduos de um determinado espaço cultural, o que implica dizer também que há uma tendência à transmissão dos conteúdos produzidos em uma cultura para outras culturas distintas. Consideramos um movimento dialógico entre as culturas e seus modos de organização, tanto das linguagens veiculadas como do conhecimento produzido.

Os sistemas sígnicos produzidos num dado contexto são imbuídos de elementos simbólicos, códigos culturais e linguagens os quais provocam o diálogo entre as comunidades e suas formas de organização. Em outras palavras, exigem estabelecer uma representação da cultura pela linguagem, além de atribuir sentido à cultura. No diálogo entre comunidades distintas é possível traçar significados atentando sempre para o fato de que a cultura é

¹⁴⁸ Segundo Machado (2003, p. 158) a concepção de estruturalidade se refere ao “Dinamismo modelizante que garante a organização de um sistema semiótico como linguagem, ainda que não possua uma língua, ou seja, uma estrutura regulada por um código definido”.

heterogênea e por isso, devemos entender que os estudos da Semiótica da Cultura nos propõe uma outra análise sobre a cultura, do ponto de vista do texto, mais especificamente do texto cultural, assim como de suas diversificadas formas de expressão. Desse modo, a Semiótica da Cultura estaria aberta à análises não somente da cultura, mas também dos possíveis sistemas complexos de signos. A cultura é um texto aberto e carrega em si sistemas sógnicos diversos que dialogam e se transformam constantemente.

4 PROCEDIMENTOS METODOLÓGICOS

A pesquisa é bibliográfica (DIETRICH, 1999; GIL, 2009, p. 50; LAKATOS; MARCONI, 2008, p. 185), pois reconhecemos que tal abordagem nos permite ter condições de conhecer a literatura com foco no tema da pesquisa, permitindo desse modo a identificação e a consulta de diversos materiais com enfoque ao tema da pesquisa. Sua natureza é qualitativa, pois levamos em consideração o tipo de interpretação dos dados que como diz Gil (2009, p. 178) “o pesquisador precisa ir além da leitura dos dados, com vistas a integrá-los num universo mais amplo em que poderão ter algum sentido”. Delineamos uma pesquisa exploratória que em acordo com Gil (2009, p. 27) possui a finalidade de esclarecer sobre um assunto ainda pouco abordado ou até mesmo desconhecido.

Para a de análise dos dados utilizamos a análise de conteúdo. O método de análise de conteúdo consiste, segundo Bardin (2011, p. 48), em:

Um conjunto de técnicas de análise das comunicações visando obter por procedimentos sistemáticos e objetivos de descrição do conteúdo das mensagens indicadores (quantitativos ou não) que permitam a inferência de conhecimentos relativos às condições de produção/recepção (variáveis inferidas) dessas mensagens.

A análise de conteúdo¹⁴⁹ enquanto método ou conjunto de técnicas permite uma análise diversificada sobre o conteúdo dos materiais. Esta técnica pressupõe processos de inferência sobre o conteúdo analisado seja por um aspecto quantitativo ou não, e dessa maneira é responsável pela representação do material analisado e ainda destaca o contexto de produção do seu conteúdo. A análise de conteúdo objetiva tanto a superação de incertezas, tendo em vista o julgamento feito sobre o conteúdo analisado, assim como da visão de quem está executando a análise. Outro objetivo é propor o enriquecimento da leitura, isto é, ir além

¹⁴⁹ A análise de conteúdo objetiva tanto a superação de incertezas, tendo em vista o julgamento feito sobre o conteúdo analisado assim como da visão de quem está executando a análise. Um outro objetivo é propor o enriquecimento da leitura, isto é, ir além das aparências do que está exposto de imediato no conteúdo dos materiais escolhidos para análise.

das aparências do que está exposto de imediato no conteúdo dos materiais escolhidos para análise.

A utilização da análise de conteúdo nesta pesquisa expõe o nosso olhar enquanto sujeitos pesquisadores sobre as diversas compreensões de linguagem e cultura no contexto tanto da Organização do Conhecimento, como na Semiótica da Cultura. No entanto, não desconsidera demais olhares e deixa em aberto para demais processos de inferência e interpretações.

No contexto desta pesquisa consideramos o método de análise de conteúdo importante por possibilitar uma análise criteriosa do material selecionado, além de pressupor várias interpretações sobre o conteúdo analisado. Nesse sentido, e com base nos objetivos da pesquisa entendemos que a análise de conteúdo é o método adequado à pesquisa.

O corpus da pesquisa compõe artigos que abarcam o campo da Organização do Conhecimento e a Semiótica da Cultura. Para isso traçamos alguns critérios. Os critérios foram: *tema*, *língua* e *acesso ao material*. O primeiro se refere ao *tema*, isto é, priorizamos pelos materiais que tivessem relação com a temática da pesquisa. Realizamos a leitura dos textos nos quais destacamos os temas como: ética transcultural, hospitalidade/garantia cultural e multilinguismo (temas presentes no núcleo de abordagens socioculturais da Organização do Conhecimento); linguagem, cultura, texto, contexto, tradução e sistemas modelizantes (temas discutidos na Semiótica da Cultura).

O segundo critério se refere à *língua*, revela que diante das nossas limitações quanto à compreensão e domínio em demais idiomas foi possível nesta pesquisa apenas a leitura de materiais na língua portuguesa, inglesa e espanhola. Por último, o critério *acesso ao material*, tendo em vista que para esse acesso é necessário a disponibilidade dos materiais.

O *corpus* teórico desta pesquisa se constitui de dois grupos: o primeiro possui trabalhos relacionados às abordagens socioculturais da Organização do Conhecimento, tendo como base os artigos disponíveis nos volumes do periódico “Advances in Knowledge Organization”¹⁵⁰ e do “Knowledge Organization (Journal)”¹⁵¹. A Knowledge Organization é o

¹⁵⁰ O período compreendido foi de 2000 a 2012, com os seguintes volumes com artigos selecionados para análise: v.8, 2002; v.10, 2006; v. 11, 2008 e v. 12, 2010. Ressaltamos ainda que não tivemos acesso aos volumes de 1 a 7.

¹⁵¹ O período compreendido foi de 1993 a 2013, com os seguintes volumes com artigos selecionados para análise: v.20, n. 4, 1993; v.24, n. 2, 1997; v.40, n.2, 1997; v.30, n.2, 2003; v.31, n.1, 2004; v.31, n.3, 2004; v.33, n.1, 2006; v.35, n. 2/3, 2008; v.27, n. 3, 2000 e v.37, n.1, 2010. É importante destacar que tivemos acesso livre aos artigos publicados a partir do ano 2000, por isto

jornal oficial bimestral da ISKO. Foi fundada em 1973 pela Dra. Ingetraut Dahlberg. Tal revista iniciou suas publicações em 1974 com o título *Classificação Internacional*. Em 1989, tornou-se o órgão oficial da ISKO e desde 1998 está sendo publicada por Ergon Verlag de Würzburg. O segundo grupo foi formado pelos artigos do periódico “*Sign Systems Studies*”¹⁵² uma revista internacional de semiótica e processos sógnicos sobre cultura e natureza, criada em 1964 por Juri Lotman.

A escolha por este corpus é justificada pelo mesmo objetivo de contemplar a temática e ainda por reunir materiais considerados significantes tanto na Organização do Conhecimento, como na Semiótica da Cultura e dessa maneira apresentam discussões e reflexões de caráter socioculturais. Realizamos a leitura dos resumos (abstracts) e das palavras-chave (keywords) dos textos relacionados ao tema da pesquisa para definir os artigos selecionados. Reunimos um total de 27 (vinte e sete) artigos pertencentes aos três periódicos. Após a análise dos textos chegamos a um número de 7 (sete) artigos relevantes para a pesquisa¹⁵³. Assim, estabelecemos categorias de análise para cada um dos grupos de materiais correspondentes ao *corpus*. As categorias que seguem no quadro abaixo foram determinadas com base em dois grupos apresentados no quadro a seguir.

foi dado ênfase aos materiais compreendidos neste período com a inclusão de outros artigos fundamentais, os quais foram publicados entre 1993 e 1999.

¹⁵² O período compreendido no levantamento foi de 1998 a 2012, destacando os seguintes volumes com artigos selecionados: v.27, 1999; v.29, n.2, 2001; v30, n.1, 2002; v. 30, n. 2, 2002; v. 33, n. 1, 2005; v.39, n. 1, 2011; v. 39, n. 2/4, 2011 e v.41, n. 2/3, 2013. Devemos informar que os volumes anteriores a este período não possuem acesso aberto, sendo somente na versão impressa, por isso não foram relacionados no levantamento. Também destacamos novamente que não incluímos os artigos em russo.

¹⁵³ Ressaltamos que também caberia utilizarmos de uma pesquisa censitária, tendo em vista que toda a população cabe em uma pesquisa exploratória. Ver BARBETTA, Pedro A. **Estatística aplicada às ciências sociais**. 7. ed. Rev. Florianópolis: Ed. da UFSC, 2007.

QUADRO 1- Descrição das categorias de análise

GRUPO 1 Abordagens Socioculturais da Organização do Conhecimento	GRUPO 2 Semiótica da Cultura
Ética transcultural	Linguagem
Transculturalismo	Cultura
Garantia cultural	Texto
Hospitalidade cultural	Contexto
Multilinguismo	Tradução
	Sistemas modelizantes

Fonte: Elaborado pela autora.

5 RESULTADOS

O levantamento feito auxiliou na configuração da análise por dispor de materiais cujo conteúdo apresentou aproximações com a discussão proposta neste estudo. Os conceitos escolhidos tanto das abordagens socioculturais como da Semiótica da Cultura se justificam pelo foco das discussões, isto é, destacam aspectos culturais e linguísticos que consideramos necessários para a organização e representação do conhecimento.

Este estudo busca contribuir com a compreensão de algumas concepções como, por exemplo, de linguagem, cultura, texto, contexto e tradução, na intenção de fornecer possíveis diálogos com a Organização do Conhecimento, mais precisamente com as discussões que envolvem as abordagens socioculturais desse campo.

Considerando as relações entre algumas perspectivas culturais defendidas nesses dois campos de estudo objetivou-se apresentar aproximações e diálogos entre as abordagens socioculturais do campo da Organização do Conhecimento e a Semiótica da Cultura. Buscamos ênfase nas concepções de linguagem, cultura, texto, contexto, tradução com destaque na denominação “tradução da tradição” e sistemas modelizantes como elementos norteadores para que pudéssemos verificar as aproximações entre as duas linhas interpretativas.

Presumimos que no domínio das discussões traçadas pela Semiótica da Cultura há elementos como a concepção de linguagem, cultura, tradução, texto, contexto, os quais são visualizados como elementos com potencialidade de diálogos, isto é, podem ser agregados e

favorecer em algum nível, teórico ou conceitual, aproximações com o campo da Organização do Conhecimento.

Buscamos resumir alguns diálogos entre a Semiótica da Cultura e as investigações socioculturais da Organização do Conhecimento. Vale ressaltar que o diálogo traçado está evidenciado pela temática de abordagem entre os dois campos de investigação. Assim, os conceitos de linguagem, cultura, texto, contexto, tradução e sistemas modelizantes encontram-se próximos tematicamente das discussões de ética transcultural, transculturalismo, garantia cultural, hospitalidade cultural e multilinguismo.

O QUADRO abaixo tem o propósito de destacar a relação entre os conceitos das abordagens socioculturais da Organização do Conhecimento e as concepções da Semiótica da Cultura.

QUADRO 2 – Diálogos teórico-conceituais

	Abordagens socioculturais da Organização do Conhecimento	Elementos da Semiótica da Cultura	Contribuições
C O N C E I T O S	Ética transcultural	Cultura	O conceito de cultura revela uma abordagem cultural ampla que dialoga com a abordagem transcultural, pois esta permite analisar a cultura tanto em seu aspecto externo e interno das comunidades e desse modo enfatiza a dimensão simbólica das comunidades.
	Transculturalismo	Cultura	A noção de cultura no contexto do transculturalismo resgata sobre a diversidade existente nos contextos das comunidades e desse modo provoca reflexões que são guiadas sobre uma crítica à concepção de multiculturalismo. O transculturalismo reflete sobre uma representação da cultura que segue do local para o global.
	Garantia cultural	Tradução, Linguagem, Texto, Contexto	Permitir um acesso à informação implica em processos de interpretação, isto é, de tradução e por isso estes devem considerar, tanto a diversidade de linguagens, como os textos produzidos nas comunidades os quais revelam os contextos culturais.
	Hospitalidade cultural	Tradução, Linguagem, Texto, Contexto	A recepção de informações implica o reconhecimento de uma diversidade cultural e linguística e por isso processos de tradução e a representação da cultura por meio de uma diversidade de linguagens podem contribuir.

	Multilinguismo	Tradução, Linguagem, Sistemas modelizantes	Frente à diversidade linguística, os instrumentos de representação como, por exemplo, os tesouros necessitam de processos de tradução que busquem reconhecer as linguagens desenvolvidas pelas comunidades, pois é preciso garantir o acesso à informação em diversas estruturas que ultrapassem as verbais, como a língua, e que permitam por meio das estruturas não verbais como as linguagens, acesso e uso da informação.
--	----------------	---	--

Fonte: Elaborado pela autora

Os diálogos evidenciados neste estudo se aproximam de uma perspectiva temática em um nível de análise interdisciplinar. Entendemos que os conceitos tratados pelas abordagens socioculturais destacam reflexões em torno de aspectos culturais que merecem atenção no tratamento da representação do conhecimento, assim como dos textos na Semiótica da Cultura. Constatamos que os diálogos ocorreram por uma relação temática e não necessariamente conceitual¹⁵⁴. Portanto, o quadro destaca uma aproximação de ordem temática cujo foco das abordagens evidencia-se pelo tratamento sociocultural das investigações e isso nos faz refletir sobre contribuições, de caráter interdisciplinar, da Semiótica da Cultura para a Organização do Conhecimento.

6 CONCLUSÕES

O conhecimento é por nós compreendido como um constructo social, e posto em uma dinâmica cultural torna-se um produto da cultura. Assim, entendemos que todo conhecimento é dinâmico e dialógico e por isso sua organização e representação necessita a compreensão da cultura e dos contextos em que esta faz parte. Organizar e representar um conhecimento são atividades de tradução da cultura e por isso é necessário considerar a diversidade linguística e cultural, e dos contextos onde estes conhecimentos são produzidos. Em outras palavras, entender a origem de qualquer conhecimento implica uma atividade semiótica. Conhecer os contextos é conhecer as culturas, as linguagens. É ainda ter o conhecimento dos elementos como códigos, símbolos e linguagens que não somente caracterizam, mas também identificam os indivíduos e as comunidades.

¹⁵⁴ Propositamente pusemos os conceitos das abordagens socioculturais paralelos com as concepções da Semiótica da Cultura, pois presumimos que haja uma relação que ainda não se define em termos conceituais, mas que caminha para uma análise interdisciplinar.

Com base na análise feita consideramos que a Semiótica da Cultura pode fornecer um diálogo que não se limita a uma dimensão conceitual, mas também teórica. Para tanto, o diálogo evidenciado neste estudo pode possibilitar uma análise interdisciplinar sobre as concepções destacadas pela Semiótica da Cultura, tendo em vista a perspectiva sociocultural apresentada pelo núcleo de abordagens socioculturais da Organização do Conhecimento.

Desse modo, relacionamos as principais contribuições da Semiótica da Cultura à Organização do Conhecimento. No entanto, acreditamos que o desdobramento desta pesquisa poderá em outras oportunidades identificar e destacar demais contribuições.

1º A concepção de cultura destaca a dimensão simbólica existente no contexto das comunidades. O conceito de cultura sugere um tratamento igualitário dos sistemas de signos. Nesse sentido, no momento de elaboração de tesouros adaptados à comunidade discursiva, os especialistas devem considerar outros sistemas de signos que não estão limitados aos signos linguísticos para propor soluções em processos de organização do conhecimento;

2º O conceito de tradução se estende a um processo interpretativo que ultrapasse a tradução de palavras no momento da transmissão dos conteúdos inseridos no sistema de representação, a tradução implica uma atividade interpretativa que considera a interpretação das comunidades de seus contextos. É necessário destacar que deve haver para os processos de tradução, a clareza, por parte dos profissionais, dos contextos e das comunidades.

3º A concepção de linguagem enfatiza a precisão sobre os processos de mediação, tendo em vista linguagens desenvolvidas nos contextos das comunidades. Assim, é preciso entender que são muitas as formas de mediação das linguagens. Desse modo, a inserção da variedade de linguagens nos processos de mediação torna complexa a representação. Por isso, a linguagem necessita ser representada para que no ato da representação os sentidos e as características das comunidades não sejam alterados, para que não comprometa e dificulte na busca da informação por essas comunidades.

REFERÊNCIAS

ALMEIDA, C. C. Sobre o pensamento de Peirce e a organização da informação. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, XI, 2010, b Rio de Janeiro. Anais... Rio de Janeiro: PPGCI/UNIRIO, 2010. Disponível em:<<http://revista.ibict.br/liinc/index.php/liinc/article/view/405/265>>

ASSIS, J.; MOURA, M. A. Social networks indexing languages and organization of knowledge: a semiotic approach. En: GNOLI, C.; MAZZOCCHI, F. (Ed.). **Paradigms and conceptual systems in knowledge organization**. Advances in Knowledge Organization Italy: ERGON VERLAG. v.12. p. 291-298. 2010.

- BARÁT, Ágnes Hajdu. Knowledge Organization in the Cross-cultural and Multicultural Society. En: ARSENAUT, C.; TENNIS, J. T. (Ed.). **Cultural and identity in knowledge organization**. Advances in Knowledge Organization. Canada: ERGON-Verlag, v. 11. p. 84-90. agosto, 2008.
- BARBETTA, Pedro A. **Estatística aplicada às ciências sociais**. 7. ed. rev. Florianópolis: Ed. da UFSC, 2007.
- BARDIN, L. **Análise de conteúdo**. São Paulo: Edições 70, 2011.
- BARITÉ, M. Organización del conocimiento: un nuevo marco teórico-conceptual en Bibliotecología y Documentación. In: CARRARA, K. (Org.). Educação, universidade e pesquisa. Marília: Unesp-Marília- Publicações, 2001. p.35-60.
- BEGHTOL, C. Domain analysis, literary warrant, and consensus: the case of fiction studies, **Journal of the American Society for Information Science and Technology**, 46, n. 1, p. 30-44, 1995.
- BEGHTOL, C. Ethical decision-making for knowledge representation and organization systems for global use. **Journal of the American Society for Information Science and Technology**, New York, v. 56, n. 9, p. 903-912, 2005.
- BEGHTOL, C. Relationships in Classificatory structure and meaning. In C.A. Bean & R. Green, (eds.). **Relationships in the organization knowledge**. Dordrecht, Netherlands: Kluwer, 2001, p. 99-113.
- BEGHTOL, C. Semantic Validity: concepts of warrant in bibliographic classification systems. **Library Resources & Technical Services**. v. 30, n. 2, p. 109-125, 1986.
- BEGHTOL, C. Universal concepts, cultural warrant, and cultural hospitality. **Advances in Knowledge Organization**. Canada: ERGON-Verlag, v.8, p. 45-49. 2002.
- BRAVO, Blanca Rodríguez. The Visibility of Women in Indexing Languages. En: Knowledge Organization for a Global Learning Society. BUDIN, G.; SWERTZ, C.; MITGUTSCH, K. (Ed.). Advances in Knowledge Organization. Austria: ERGON-Verlag, v. 10. p. 413-422. 2006.
- CAFÉ, L. M. A.; BRÄSCHER, B. M. Organização do conhecimento: teorias semânticas como base para estudos e representação de conceitos. *Inf. Inf.*, Londrina, v.16. n.3. p. 25-51, jan. / jun.2011.
- CAMPBELL, G. Queer theory and the creation of contextual subject access tools for gay and lesbian communities. *Knowledge Organization*, v.27, n. 3, 2000. p. 122-131.
- CAMPBELL, G. Tensions between language and discourse in north american knowledge organization, *Knowledge Organization*, v. 37, n. 1, 2010. p. 51-57.
- CAMPOS, M. L. de A. Modelização de domínios de conhecimento: uma investigação de princípios fundamentais. *Ciência da Informação*, Brasília, v. 33, n. 1, p. 22-32, 2004.
- CANCLINI, N. G. **Culturas híbridas: estratégias para entrar e sair da modernidade**. São Paulo: Edusp, 2008.

CINTRA, A. M. et. al. **Para entender as linguagens documentárias**. 2. ed. rev. ampl., São Paulo: Polis, 2002.

DAHLBERG, I. Knowledge organization: a new science? **Knowledge Organization**, Frankfurt, v.33, n.1. p.11-19. 2006.

DAHLBERG, I. Knowledge organization: its scopes and possibilities. **Knowledge Organization**. Frankfurt, v. 20, n. 4, p. 211-222, 1993.

DAL' EVEDOVE, P. R. O tratamento temático da informação em abordagem sociocultural: diretrizes para definição de política de indexação em bibliotecas universitárias. 2014. 259f. Tese (Doutorado em Ciência da Informação) Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2014.

DIETERICH, H. **Nueva, guía para la investigación científica**. México: Planeta, 2001. 237p.

ESTEBAN NAVARRO, M. A. El marco disciplinar de los lenguajes documentales: la Organización del Conocimiento y las ciencias sociales. **Scire**, v. 2, n. 1, p. 93-107. 1996.

FUJITA, M. S. L. ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO NO BRASIL: análise de aspectos conceituais e da produção científica do ENANCIB no período de 2005 a 2007. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 1, n. 1 2008. Disponível em: <<http://inseer.ibict.br/ancib/index.php/tpbci/article/viewArticle/4>>

GARCÍA GUTIÉRREZ, A. Cientificamente favelados: uma visão crítica do conhecimento a partir da epistemografia. **Transinformação**, Campinas, v.18, n.2, p.103-112, maio/ago. 2006.

GARCÍA GUTIÉRREZ, A. Declassification in knowledge organization: a post-epistemological essay. **Transinformação**, Campinas, v.23, n. 1, p. 5-14, jan. /abr., 2011.

GARCÍA GUTIÉRREZ, A. Exomemoria y cultura de frontera: hacia una ética transcultural de la mediación. VII Congreso Internacional sobre Organización del Conocimiento celebrado en Granada (España), 2002c.

GARCÍA GUTIÉRREZ, A. Knowledge organization from a “culture of the border”: towards a transcultural ethics of mediation. In: LÓPEZ-HUERTAS, M. J. (eds.). *Challenges in knowledge representation and organization for the 21st century: integration of knowledge across boundaries*. Würzburg: ERGON-Verlag, v.8, pp. 516-522.2002a.

GARCÍA GUTIÉRREZ, A. La memoria subrogada: mediación, cultura y conciencia en la red digital. Granada: Editorial de la Universidad de Granada, 2002b. Arquivo em meio magnético.

GARCÍA GUTIÉRREZ, A. Otra memoria es posible: estrategias descolonizadoras del archivo mundial. Sevilla: Universidad de Sevilla, 2004.

GARCÍA GUTIÉRREZ, A. **Outra memória é possível**: estratégias descolonizadoras do arquivo mundial. Petrópolis: Vozes, 2008.

GARCÍA GUTIÉRREZ, A. Principios de lenguaje epistemográfico: la representación del conocimiento sobre Patrimonio Histórico Andaluz. Sevilla: Instituto Andaluz del Patrimonio Histórico, 1998. (Cuadernos técnicos, 3).

GARCÍA GUTIÉRREZ, A. Proyectar la memoria: del ordo nacional a la reapropiación crítica. *Transinformação*, Campinas, v. 15, n. 1, p. 7-13, 2003.

GIL, A. C. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Editora Atlas S.A., 2009.

GREIMAS, A. J.; COURTÉS, J. **Dicionário de semiótica**. São Paulo: Editora Contexto, 2008.

GUIMARÃES, J. A. C. Aspectos éticos em organização e representação do conhecimento (O.R.C): uma reflexão preliminar. Marília, 2005. 24p. Conferência apresentada na mesa-redonda: Organização do conhecimento e interdisciplinaridade. In: Seminário: Memória, informação e organização do conhecimento: cruzando fronteiras da identidade. Rio de Janeiro, Unirio, 25 e 26 de agosto de 2005.

GUIMARÃES, J. A. C.; MILANI, S. O. Bias in the indexing languages: theoretical approaches about feminine issues. En: GNOLI, C.; MAZZOCCHI, F. (Ed.). **Paradigms and conceptual systems in knowledge organization**. Advances in Knowledge Organization Italy: ERGON VERLAG. v.12. p. 424-429. 2010.

GUIMARÃES, J. A. C.; MILANI, S. O. Problemas éticos em representação do conhecimento: uma abordagem teórica. **DataGramZero**, v.12, n.1, fev./2011.

GUIMARÃES, J. A. C.; PANDO, D. A. O Profissional da informação e as atividades de organização da informação / conhecimento. **Interatividade**, v.1, n. 2, 2006.

GUIMARÃES, J. A. C.; PINHO, F. A. Reflexiones acerca de las teorías sobre ética en la representación del conocimiento. *Biblios. Revista de bibliotecología y Ciencias de la Información*. Jul. /set., 2008. pp. 1-12.

GUIMARÃES, J.A.C.; DOBEDEI, V. (Orgs.). Desafios e perspectivas científicas para a organização e representação do conhecimento na atualidade. [recurso eletrônico]. Marília: ISKO- Brasil: FUNDEPE, 2012.

HALLIDAY, M. A. K. **El Lenguaje como semiótica social**: la interpretación social del lenguaje y del significado. México: Fondo de Cultura Económica, 1982. 327p.

HJORLAND, B. Concepts, paradigms and knowledge organization. En: GNOLI, C.; MAZZOCCHI, F. (Ed.). **Paradigms and conceptual systems in knowledge organization**. Advances in Knowledge Organization. Italy: ERGON VERLAG. v.12. p. 38-42. 2010.

HJORLAND, B. Fundaments of knowledge organization. **Knowledge Organization**, v.30, n. 2, p.87-111, 2003.

HJORLAND, B. Semantics and Knowledge organization. **Annual Review of Information Science and Technology**, v.41, n.1,p. 367-405. 2007. Disponível em:<http://www.academia.edu/1912560/Semantics_and_knowledge_organization>

HJORLAND, B. Theories of knowledge organization-Theories of knowledge. **Knowledge Organization**, v. 40, n. 3, p 169-181. 2013.

HUDON, M. Multilingual thesaurus construction: integrating the views of different cultures in one gateway to knowledge and concepts. **Knowledge Organization**. v. 24, n.2, p.84-91. 1997.

- HUDON, M. True and tested products: thesauri on the web. **The Indexer**, London, v. 23, n. 3, p. 115-119, 2003. ISKO. About ISKO. Disponível em: <<http://www.isko.org/about.html>>
- HUDON, M. HUDON, M. Accessing documents and information in a world without frontiers. **The Indexer**, London, v. 21, n. 4, p. 156-159, 1999.
- JAKOBSON, R. **Linguística e Comunicação**. São Paulo: Cultrix. 1995.
- LAKATOS, E. M.; MARCONI, M. de A. **Fundamentos de metodologia científica**. 6.ed. São Paulo: Atlas, 2008.
- LÓPEZ-HUERTAS, M. J. Análisis del dominio interdisciplinar para la representación y organización del conocimiento. En: GONZÁLEZ DE GÓMEZ, M. N.; ORRICO, E. G. D. **Políticas de memória e informação: reflexões na organização do conhecimento**. Natal: EDUFRN, p. 209-235. 2006
- LÓPEZ-HUERTAS, M. J. Cultural impact on Knowledge Representation and Organization in a Subject Domain. En: ARSENAUT, C.; TENNIS, J. T. (Ed.). **Cultural and identity in knowledge organization**. Advances in Knowledge Organization. Würzburg: ERGON VERLAG. v.11. p. 304-346. 2008.
- LÓPEZ-HUERTAS, M. J. Epistemological dynamics in scientific domains and their influence in knowledge Organization. En: GNOLI, C.; MAZZOCCHI, F. (Ed.). **Paradigms and conceptual systems in knowledge organization**. Advances in Knowledge Organization Italy: ERGON VERLAG. v.12. p. 91-97. 2010.
- LOTMAN, I. M. El símbolo en el sistema de la cultura. **Entretextos. Revista Electrónica Semestral de Estudios Semióticos de la Cultura**. n. 2, Nov. 2003b. Disponível em: <<http://www.ugr.es/~mcaceres/entretextos/pdf/entre2/escritos/escritos4.pdf>>
- LOTMAN, Y. M. **Universe of the mind: a semiotic theory of culture**. London: I.B. Tauris & CO. LTD, 1990.
- LOTMAN, Y. M. **Cultura y explosión: lo previsible y lo imprevisible en los procesos de cambio social**. España: Gedisa editorial, 1999.
- LOTMAN, Y. M. Sobre el concepto contemporáneo de texto. Entretextos. **Revista Electrónica Semestral de Estudios Semióticos de la Cultura**, (2). nov. 2003^a. Disponível em: <<http://www.ugr.es/~mcaceres/entretextos/pdf/entre2/lotman.pdf>>
- MACHADO, I. A comunicação de sínteses em perspectiva semiótico-evolutiva: modelização de linguagens, de formato, do signo informativo, de temporalidades. **Revista Fronteiras, estudos midiáticos**. v. 12, n. 2, p. 95-104, maio / ago. 2010.
- MACHADO, I. **Escola de Semiótica: a experiência de Tártu - Moscou para o estudo da cultura**. São Paulo: Ateliê Editorial, 2003. 189p.
- MACHADO, I. **Semiótica da Cultura e Semiosfera**. São Paulo: Annablume / Fapesp, 2007.
- MAI, J. -E. Classification in context: relativity, reality, and representation. **Knowledge Organization**. v.31, n. 1, p. 39-48. 2004.
- MAI, J-E. Semiotics and indexing: an analysis of the subject indexing process. **Journal of Documentation**, London, v.57, n. 5, p. 591-522, sep. 2001.

MAI, J-E. The concept of subject in a semiotic light. In: SCHWARTS, C.; RORVIG, M. (ed.). **Digital collections: implications for users, funders, developers and maintainers.** Medford, NJ: Information Today, p. 54-64.1997a..(Proceedings of the ASIS Annual Meeting; 34).

MAI, J-E. The concept of subject: on problems in indexing. In: McILWAINE, I. C. (ed.). **Knowledge organization for information retrieval: 6th International Study Conference on Classification Research.** The Hague: FID, 1997b. p. 60-67. (FID, n. 716).

MILANI, S. O. **Estudos éticos em representação do conhecimento: uma análise da questão feminina em linguagens documentais brasileiras.** 2010. 141 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010.

MONTEIRO, S. D. O ciberespaço e os mecanismos de busca: novas máquinas semióticas. *Ciência da Informação*, Brasília, v. 35, n. 1, p. 31-38, jan./abr. 2006b.

MONTEIRO, S. D. Semiótica peirceana e a questão da informação e do conhecimento. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, Florianópolis, 2o n. especial, p. 43-57, 2º sem. 2006a. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb>> Acesso em: 15 out. 2006.

MORRIS, C. W. **Fundamentos da teoria do signo.** Rio de Janeiro: Eldorado; São Paulo: Universidade de São Paulo, 1976.

MOURA, M. A. Ciência da informação e semiótica: conexão de saberes. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, 2o n. especial, p. 1-17, 2o sem. 2006. Disponível em:< <http://www.periodicos.ufsc.br/index.php/eb>>. Acesso em: 15 out. 2006b.

MOURA, M. A. Interoperabilidade semântica e ontologia semiótica: a construção e o compartilhamento de conceitos científicos em ambientes colaborativos online. **Inf.Inf.**, Londrina, v.16. n. 3. p. 165-179, jan./jun. 2011.

MOURA, M. A. Leitor-Bibliotecário: interpretação, memória e as contradições da subjetividade em processos de representação informacional. In: KURAMOTO, H.; NAVES, M. M. L. (Orgs.). **Organização da informação: princípios e tendências.** Brasília: Brinquet de Lemos/Livros, 2006a. p 22-35.

MOURA, M. A. Signi-fica ou signi-vai? As teorias da significação no campo da Ciência da Informação. In: REIS, A. S.; CANRAL, A. M. (Org.) **Informação, cultura e sociedade: interlocuções e perspectivas.** Belo Horizonte: Novatus, 2007, p. 61-80.

MOURA, M. A.; SILVA, A. P.; AMORIM, V. R. A concepção e o uso das linguagens de indexação face às contribuições da Semiótica e da Semiologia. **Informação & Sociedade: Estudos**, João Pessoa, PB, v. 12, n.1, p.1-22, 2002. Disponível em : <<http://www.informacaoesociedade.ufpb.br/issuev12n102.htm>>. Acesso em: 13 out. 2005.

NÖTH, W. **A semiótica do século XX.** São Paulo: Annablume, 1996.

PEIRCE, C. S. **Semiótica.** 2.ed. São Paulo: Perspectiva, 1990.

PINHO, F. A. **Aspectos éticos em representação do conhecimento em temáticas relativas à homossexualidade masculina: uma análise da precisão em linguagens de indexação**

brasileiras. 2010. 149 f. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2010.

PINHO, F. A. **Aspectos éticos em representação do conhecimento**: em busca do diálogo entre Antonio García Gutiérrez, Michèle Hudon e Clare Beghtol. 2006. 123 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2006.

SAUSSURE, F. de. **Curso de lingüística geral**. 2. ed. São Paulo: Cultrix, 1970

SCHNAIDERMAN, B. **Semiótica Russa**. São Paulo: Perspectiva, 1979.

SEMPRINI, A. **Multiculturalismo**. Bauru, SP: EDUSC, 1999. 178 p.

SMIRAGLIA, R. P. The progress of theory in knowledge organization. *Library Trends*, v. 50, n. 3, p. 330-349, 2002.

SOUSA, B. P. de. ; ALMEIDA, C. C. Um olhar semiótico sobre o processo de indexação: a questão da representação e do referente. *Inf. & Soc.: Est., João Pessoa*, v.22, n. 2, p. 23-34, maio / ago. 2012.

THELLEFSEN, T. L. **Fundamental signs and significance-effects**: a semeiotic outline of fundamental signs, significance-effects, knowledge profiling and their use in knowledge organization and branding. 2009. 284 f. Doctoral Dissertation – Department of international Culture and Communication Studies. Copenhagen Business School.

THELLEFSEN, T. L. Knowledge profiling: the basis for knowledge organization. **Library Trends**, v. 52, n. 3, p. 507-514, winter 2004b.

THELLEFSEN, T. L. **Pragmaticism and the role of terminology**. *Impact: an electronic journal on formalisation in text, media and language*, abr. 2003. Disponível em: <<http://www.impact.hum.auc.dk>>. Acesso em: 05 maio 2007.

THELLEFSEN, T. L. Semiotic knowledge organization: theory and method development. **Semiotica**, v. 142, n. 1 / 4, p. 71-90, 2002.

THELLEFSEN, T. L. The fundamental sign. **Semiotica**, v. 149, n.1/4, p. 245-259, 2004a.

THELLEFSEN, T. L.; THELLEFSEN, M. M. Pragmatic semiotics and knowledge organization. **Knowledge Organization**, Würzburg, v. 31, n. 3, p. 177-187, 2004.

TERMINOLOGIA E DOCUMENTAÇÃO: ESTUDO TERMINOGRÁFICO DA PERFORMANCE MUSICAL

TERMINOLOGY AND DOCUMENTATION: a terminology study on musical performance

Soraia de Andrade Lara Carvalho
Gercina Ângela B. de Oliveira Lima

Resumo: Aborda os fundamentos da Terminologia e as suas contribuições na definição de campos conceituais da área de Música. Adotaram-se as técnicas da terminografia para a realização de estudo terminológico pontual do termo *performance* musical. O *corpus* foi constituído por artigos publicados em três periódicos científicos da área de Música e de trabalhos apresentados no XXII Encontro da Associação Nacional de Pesquisa e Pós-Graduação em Música, 2012. A análise dos dados revelou a presença de um repertório de termos adjacentes ao termo estudado, considerados significativos e que, categorizados em assuntos específicos, constituíram um conjunto terminológico homogêneo e representativo da área. Identificou-se a variação terminológica e estabeleceram-se as relações conceituais entre os termos que foram organizados a partir das linguagens de especialidade, demonstrando que, quando se opera com as palavras em funcionamento, é possível atingir a desejada contiguidade e semelhança entre o vocabulário utilizado na representação da informação e o universo do conhecimento expresso nos documentos.

Palavras-chave: Teoria da Terminologia. Terminologia musical. Linguagens de especialidade. Linguagens documentárias.

Abstract: The research presented here sought to study the fundamentals of Terminology and their contributions in the definition of conceptual fields from specialized languages present in the scientific discourses of Music area. A study was conducted terminological points with the term musical performance items located in three journals Music and the papers presented at the Anais da Associação Nacional de Pesquisa e Pós-Graduação em Música, all published in 2012. The data analysis showed that there is a meaningful collection of terms adjacent to the term musical performance that can be categorized as specific subjects to constitute an homogeneous and representative vocabulary for the area. It was identified a terminological variation and established the conceptual relations between the terms, showing the different ways of treatment of the concept musical performance given in the scientific texts about the subject. The terms were organized from the languages of speciality, demonstrating that, when we work with the words used by the specialists in their texts, it is possible to get the necessary matching between the vocabulary used for information representation and the universe of knowledge expressed in documents.

Keywords: Theory of Terminology. Musical Terminology. Speciality languages. Documentary languages.

1 INTRODUÇÃO

A produção de conhecimentos está intimamente ligada à criação de novos termos, pois a todo e qualquer campo de conhecimento corresponde um conjunto de noções que lhe é próprio. Cada ideia precisa ser designada por uma unidade terminológica, a fim de ser comunicada ao público interessado, visto que o progresso do conhecimento se dá através da

transferência de informação. Acompanhando a evolução da produção técnico-científica, as linguagens de especialidade¹⁵⁵ se multiplicam e se generalizam, tornando-se uma necessidade cotidiana o seu entendimento e organização para garantir a efetiva comunicação do conhecimento especializado.

Pesquisadores empenham-se em fazer avançar a ciência e a tecnologia, divulgando os resultados de seus trabalhos nos mais variados tipos de documentos. No ambiente científico, o texto é considerado um veículo institucional que tem como objetivo operacionalizar a comunicação do conhecimento. Por meio dele, o pesquisador registra e expõe o resultado de suas atividades de pesquisa na busca pela aprovação e legitimação entre os pares, através das linguagens de especialidade nas diversas modalidades: falada, escrita, em meio físico ou eletrônico que se movimentam entre os especialistas da área. Gardin afirma que “a expressão concreta de um raciocínio científico é o texto científico, no qual o autor expõe as operações do espírito que o conduziram da observação de certos fatos empíricos ao enunciado de proposições” (Gardin, 1987 *apud* KOBASHI, 1997, p. 2).

A relação entre o texto e os termos acontece de forma complementar na constituição das linguagens de especialidade, pois é no texto especializado que as palavras adquirem o *status* de termo. As unidades terminológicas são ocorrências que podem expressar nomes próprios, coisas, locais, siglas, etc., e também os neologismos, sendo que o fortalecimento desses termos se dá, concomitantemente, com o desenvolvimento do domínio no qual estão inseridos. Para Benveniste (1989):

A constituição de uma terminologia própria marca, em toda ciência, o advento ou o desenvolvimento de uma conceitualização nova, assinalando, assim, um momento decisivo de sua história. Poder-se-ia mesmo dizer que a história particular de uma ciência se resume na de seus termos específicos. Uma ciência só começa a existir ou consegue se impor na medida em que faz existir e em que impõe seus conceitos, através de sua denominação. (BENVENISTE, 1989, p. 252)

Dominar ou ter acesso às terminologias das linguagens de especialidade torna-se hoje uma condição essencial para se garantir a comunicação do conhecimento, entretanto, a definição dos campos nocionais (conceituais) de áreas do conhecimento que estão se formando ou se transformando, geralmente é dificultada pela carência de obras referenciais que reúnem, organizam e disseminam as terminologias específicas das áreas.

¹⁵⁵ Línguas (ou linguagens) de especialidade são consideradas sistemas de comunicação oral ou escrita usados por uma comunidade de especialistas de uma área particular do conhecimento (Pavel; Nolet, 2002, p. 124).

No campo da Música, o desenvolvimento da produção científica rumo à especialização do conhecimento, concorreu para o surgimento de diversas linguagens de especialidade presentes tanto nos discursos científicos como nas demandas de informação dos usuários dos Sistemas de Recuperação da Informação (SRIs). Esta constatação decorre de observações e reflexões feitas pela autora em sua atividade profissional na Biblioteca da Escola de Música da Universidade Federal de Minas Gerais (UFMG), onde atua na organização da informação musical.

Nos anos de 1980, verificou-se a especialização no ensino e estudo da Música, alicerçada na criação de vários programas de pós-graduação e na realização de diversos eventos científicos da área, fato que delineou novos cenários para a pesquisa musical no Brasil. De acordo com Assis *et al*:

[...] a restrição aos trabalhos de análise musical e a busca por repertórios para a prática interpretativa, deram lugar ao interesse pelo conhecimento pluridisciplinar da Música, sendo que as abordagens passaram a privilegiar a Música na cultura, a relação compositor-obra-intérprete-receptor e o papel do artista em sua sociedade. (ASSIS *et al*, 2009, p. 13)

Desta forma, a prática musical deixou de ser o principal objetivo dos modelos de ensino dando lugar a novas abordagens de caráter interdisciplinar. Iniciou-se um diálogo profícuo entre a Música e outras áreas do conhecimento como a História, Sociologia, Antropologia, Psicologia, Teoria da Literatura, Linguística e Educação, que sinalizou para um momento de renovação e aprofundamento na produção científica rumo à especialização do conhecimento.

Este fato trouxe grandes desafios para o tratamento e a organização da informação musical, suscitando as seguintes reflexões acerca das práticas inerentes a estes processos: uma vez que as áreas do conhecimento em formação ou transformação são carentes de vocabulários mais estruturados devido à própria imaturidade das linguagens de especialidade, no campo da Música constatou-se a falta de instrumentos de controle terminológico, em língua portuguesa, adequados para subsidiar as atividades de representação e recuperação da informação presente em novos conteúdos informacionais advindos da especialização deste conhecimento.

No desempenho da atividade de indexação pela pesquisadora e durante a análise de termos descritores encontrados em catálogos eletrônicos de bibliotecas de instituições brasileiras de ensino de Música e no catálogo da Fundação Biblioteca Nacional, verificou-se a adoção de termos extraídos da Lista de Cabeçalhos de Assuntos da Biblioteca do Congresso Americano ou *Library of Congress Subject Headings* (LCSH).

A adoção desta lista na representação dos assuntos pode ser considerada uma barreira para o fortalecimento da linguagem de especialidade da área, uma vez que não retrata o domínio da Música. Esta Lista foi criada por especialistas da Biblioteca do Congresso Americano, a partir da descrição temática dos documentos musicais, um a um, existentes naquela Biblioteca, para atender às necessidades de representação do seu acervo. Como resultado, este produto carrega em sua estrutura de significados, interferências sociais e culturais do contexto específico para o qual foi criada não atendendo, satisfatoriamente, outras perspectivas de análise. Portanto, “não se garante a representação do sistema conceitual de um domínio, apenas com o uso dos termos de uma área - tal como foi realizado o processo de garantia literária utilizado na elaboração da *LCSH*” (LARA, 2002, p. 134).

Outro fator negativo na adoção da referida lista é o emprego de palavras traduzidas do inglês, muitas vezes com adaptações por não existir a equivalência em português, o que resulta numa interferência na identificação cultural dos seus termos, uma vez que são representações sócio-historicamente construídas.

Há que se considerar também que a informatização dos acervos ocorrida nos últimos anos acarretou várias mudanças de procedimentos tanto nas atividades do profissional da informação como no comportamento da comunidade usuária dos SRIs. Em relação ao profissional, a atividade de intercâmbio dos dados bibliográficos através da catalogação cooperativa, possibilitou uma nova rotina de trabalho em que a filosofia é de se evitar a duplicidade na descrição das informações bibliográficas e dedicar maior atenção à descrição das informações temáticas dos documentos. A diversidade da informação musical, registrada em variados suportes e formatos como recursos de mídia, imagem e som, e a rapidez em que é produzida, vêm ampliando, significativamente, o universo bibliográfico a ser disponibilizado, exigindo novas possibilidades de análise e organização dos dados.

Em relação ao usuário, a disponibilização de registros bibliográficos e documentos em formato eletrônico oferecem a ele maior autonomia em suas buscas em virtude das novas formas de identificar, localizar, acessar e obter as informações desejadas. Entretanto, o cenário que se configura de crescimento da oferta e procura de informação têm suas deficiências agravadas em relação às lacunas percebidas no processo de representação e recuperação da informação, especialmente no que diz respeito às linguagens documentárias, instrumentos de controle de vocabulário adotados na representação da informação.

Observou-se, no campo da Música, que as linguagens documentárias que deveriam atuar como instrumentos de comunicação não estão contemplando, satisfatoriamente, os termos e conceitos presentes nas linguagens de especialidade decorrentes do desenvolvimento

da pesquisa. A constatação da incompatibilidade entre a linguagem dos especialistas, linguagem dos sistemas de informação e linguagem dos usuários é um fato que prejudica o acesso à informação e ao conhecimento especializado. Desta forma, ao inserir o sistema conceitual como parte destas reflexões, esta pesquisa teve como objetivo geral estudar os fundamentos da Terminologia para demonstrar de que maneira a circunscrição pelo campo semântico de uma área de especialidade pode contribuir na definição do seu campo nocional e assegurar a efetiva comunicação entre os SRIs e a comunidade usuária.

Segundo Bélanger (1991, *apud* GALVÃO 2004, p. 246) “o conhecimento de uma língua de especialidade não significa compreender, detalhadamente, todos os conceitos que a compõe, mas sim, conhecer seus termos e conceitos mais importantes que circulam com a comunicação do conhecimento”. Para demonstrar a contribuição da Terminologia na identificação das linguagens de especialidade que transitam pelos ambientes de produção e transferência do conhecimento musical, estabeleceram-se como objetivos específicos desta pesquisa realizar o estudo terminográfico do termo *performance* musical e avaliar o emprego da terminologia na representação da informação.

2 A REPRESENTAÇÃO DA INFORMAÇÃO

Como consequência do desenvolvimento científico e tecnológico, a especialização do conhecimento musical tem gerado um grande volume de informações, que precisam ser adequadamente tratadas para viabilizar o seu registro, transferência e assimilação, contribuindo, desta maneira, para a concretização do ciclo de produção e comunicação do saber. Neste estudo, a ênfase foi dada à representação temática dos documentos que é definida por Cintra (2002, p. 34), como uma “atividade metodológica específica no interior da Documentação, que trata da análise, síntese e representação da informação, com o objetivo de recuperá-la e disseminá-la”. Esta atividade se desdobra nos procedimentos de análise conceitual - que implica em determinar o assunto do documento, através da compreensão do conteúdo temático da obra e dos objetivos do autor - e da identificação e seleção de termos válidos para a sua representação.

A escolha dos elementos de representação temática deve levar em conta a maneira como o autor expõe as ideias no texto, bem como as necessidades informacionais dos usuários potenciais de um sistema de informação, a fim de promover um processo de comunicação. Para que isto ocorra, Novellino (1996, p.37) considera que a Ciência da Informação dá ênfase ao desenvolvimento e análise de teorias, visando à elaboração de instrumentos e métodos para

a análise de assunto, construção de linguagens documentárias e geração de sistemas de classificação.

Nasce daí a relação da Ciência da Informação com a linguagem natural que, através dos sistemas de organização do conhecimento ou *Knowledge Organization Systems (KOS)* - que têm a função de comunicar o conhecimento – permitem que os profissionais da informação e usuários compartilhem um mesmo vocabulário, respectivamente, na representação e recuperação da informação. Segundo Lara, para se falar realmente em transmissão de informação e garantir que o usuário obtenha sucesso em suas buscas, as linguagens documentárias devem ser elaboradas explicitando “hipóteses de organização da informação, que sejam razoavelmente compartilhadas entre a comunidade de usuários e o sistema que se propõe” (LARA, 2002, p. 132).

Sendo construídas para organizar e facilitar o acesso e a transferência da informação, as linguagens documentárias se baseiam nas linguagens de especialidade, bem como em palavras da linguagem natural, presentes nos discursos científicos. Rumo à especialização do conhecimento, as linguagens documentárias estão sendo construídas para representar domínios cada vez mais específicos, abandonando a ideia inicial pretensiosa de cobrir todo o universo do conhecimento. Sua estruturação vai variar segundo o estágio de desenvolvimento de cada domínio do conhecimento, podendo existir até mais de uma linguagem documentária para o mesmo domínio, uma vez que refletem pontos de vista particulares sobre o conhecimento e respondem às necessidades específicas de informação.

Nesse direcionamento, a tendência à especialização das linguagens documentárias têm-se apoiado nas linguagens de especialidade utilizadas para a delimitação de campos nocionais específicos do saber. Segundo Lara (2002), os trabalhos terminológicos constituem uma ferramenta imprescindível para a construção da linguagem documentária já que permitem conferir referência aos descritores. A autora afirma que:

Se antes a biblioteconomia e a documentação trabalhavam empiricamente a partir de referências da classificação filosófica (enfatizando apenas a segmentação, mas não sua expressão linguística) e, em seguida, para dar conta das formas significantes, a partir de palavras (através dos processos de extração baseados em frequência ou ocorrência) ou da seleção empírica de unidades significativas (sem a definição daquilo que poderia caracterizar uma unidade significativa), a partir da Terminologia ela passa a contar com instrumentos que trabalham com o termo, unidade que representa o conceito dentro de um domínio ou área de atividade. (LARA, 2002, p. 136)

Portanto, a Terminologia pode oferecer uma importante contribuição à Ciência da Informação no sentido de propor uma solução aplicada para resolver problemas de delimitação de universos conceituais, quando não se tem, ainda, obras especializadas para

realizar esta estruturação. Através da identificação das linguagens de especialidade, são evidenciadas as referências socialmente validadas entre as comunidades. Dessa forma, o trabalho deixa de ser realizado de forma empírica de categorização e passa a contar com um arcabouço teórico e metodologicamente sedimentado em referências concretas dos domínios de especialidade.

3 TERMINOLOGIA

A Terminologia, na qualidade de disciplina científica que estuda as linguagens de especialidade e seus vocabulários, desempenha um papel fundamental no processo de representação e transmissão do conhecimento. Representada por um termo polissêmico, quando grafada com a inicial maiúscula, a Terminologia é uma disciplina científica que estuda as chamadas línguas (ou linguagens) de especialidade, fornecendo metodologia para a descrição, ordenamento e transferência do conhecimento e indicando princípios que regem a compilação e a formação dos termos, a estruturação de campos conceituais e a administração das unidades terminológicas. Quando grafada com a inicial minúscula, a terminologia refere-se ao conjunto de termos de uma área, relacionados e definidos rigorosamente para designar as noções que lhe são úteis, constituindo um sistema de termos organizados, a partir de noções particulares (CINTRA, 2002, p. 38).

Seus estudos têm-se desenvolvido de modo intenso nas últimas décadas acompanhando o avanço técnico-científico e a aplicação prática dos seus fundamentos têm sido objeto de projetos de trabalhos dos mais diferentes profissionais que lidam com a linguagem e a criação de meios eficientes de tratamento da informação, de modo a garantir a sua comunicação. A Terminologia constitui-se em um campo do conhecimento que, ao dialogar com diferentes áreas especializadas, capacita-se a estabelecer princípios e métodos de elaboração de ferramentas e produtos, tais como sistemas de reconhecimento automático de terminologias, glossários, dicionários técnico-científicos e bancos de dados terminológicos.

Para a Ciência da Informação, interessa a aproximação que acontece entre a Documentação e a Terminologia, tendo em vista o objetivo maior da organização da informação para a sua recuperação. Segundo Cabré, a existência de pontos de confluência entre a Documentação e a Terminologia está na relação bilateral entre ambas, pois “a terminologia é uma peça necessária ao trabalho de documentação e os documentos são imprescindíveis para o trabalho terminológico” (CABRÉ, 1999, p. 233).

Emprega-se a Terminologia (teórica e concreta) para identificar as terminologias das áreas, como subconjuntos da língua nos quais as palavras têm significados específicos, de

acordo com o contexto em que se inserem, podendo ser classificadas segundo os seus objetivos. Cabré (1993), considera que a terminologia descritiva ocupa-se com a coleta de dados e a descrição dos termos e a terminologia normativa preocupa-se com a uniformização de conceitos e atribuição de termos para designá-los.

Embora o emprego de termos técnico-científicos seja antigo, o surgimento de um campo de estudos dedicado à Terminologia é mais recente, situando-se na segunda metade do século XX, com Eugene Wüster (1898-1977). Em sua tese de doutorado - *Internationale Sprachnormung in der Technik* - Wüster abordou a necessidade da normalização da linguagem técnica, expôs os motivos que justificavam a sistematização dos métodos de trabalho em Terminologia e estabeleceu os princípios que deveriam reger os trabalhos sobre os termos. Sua preocupação era com o trabalho normativo e metodológico, o que o levou a propor uma metodologia para tratamento de dados terminológicos, baseada na sistematização dos conceitos para elaboração de dicionários terminológicos (BARROS, 2004, p. 53).

Em 1931 Wüster publicou um livro no qual elaborou uma análise detalhada da Terminologia, considerada por ele como instrumento de comunicação abordando a natureza e a descrição (definição) dos conceitos, as relações entre eles, a formação dos termos, a internacionalização dos conceitos e dos termos dentre outros tópicos, constituindo-se, desta forma, a base da Teoria Geral da Terminologia (TGT).

De caráter pragmático - para favorecer a comunicação das ciências no plano internacional - a teoria de Wüster foi objeto de estudos, reflexões e críticas. Apesar de conter uma acurada compreensão dos mecanismos dos léxicos terminológicos, a obra de Wüster não ampliou seu poder explicativo e, com isto, na última década do século vinte, verificou-se um redirecionamento dos estudos terminológicos, com o início de um novo percurso “pautado pelo incremento de investigações terminológicas de base linguístico-comunicacional” (KRIEGER; FINATTO, 2004, p. 34).

Neste cenário, os termos deixaram de ser vistos como simples palavras estáticas que figuravam nos dicionários e passaram a ser mais valorizados no aspecto comunicativo das linguagens de especialidade, em detrimento dos propósitos normalizadores. A particularidade desse novo pensamento, que é a base da Teoria Comunicativa da Terminologia (TCT), de Maria Teresa Cabré, está na proposta de compreender o termo à luz de um ponto de vista descritivo, contrapondo-se às limitações da TGT, de Wüster. Esta teoria almeja reconhecer a existência da variação conceitual e denominativa nos domínios de especialidade, é mais flexível e leva em conta a dimensão textual e discursiva dos termos, fundamentando-se numa visão mais comunicativa da linguagem.

Com base nos fundamentos da TCT que pelas características citadas acima permite incorporar melhor os referenciais das linguagens de especialidade e linguagens dos usuários e seguindo os princípios da Terminografia – que é a parte aplicada da Terminologia – explicitam-se, a seguir, os procedimentos metodológicos adotados no estudo terminológico pontual do termo *performance* musical.

4 PROCEDIMENTOS TEÓRICO-METODOLÓGICOS

Para a realização do estudo da terminologia especializada, elegeu-se o domínio da Música por ser esta a área de atuação profissional da pesquisadora no tratamento e organização da informação musical. Como subdomínio do conhecimento, selecionou-se a *Performance* Musical que, de acordo com Borém; Ray (2012) é a subárea da Música que mais se desenvolveu, nos últimos anos, no cenário da pesquisa musical brasileira.

O *corpus* da pesquisa foi constituído por trabalhos publicados no ano de 2012, em artigos de três revistas científicas do domínio da Música, classificadas no estrato A do *Sistema Qualis* da Capes: *Per Musi*, *Música Hodie* e *Opus*, e pelos *Anais do XXII Encontro da Anppom*, focalizando a subárea *Performance* Musical, totalizando 144 artigos científicos. Para a identificação de ocorrência dos termos utilizou-se o recurso “localizar e marcar”, do editor de textos *Word*, para pesquisar a palavra *performance* e a expressão *performance* musical. O recolhimento de todas as definições e contextos de uso dos termos com a transcrição literal do trecho onde foram localizados propiciou um aprofundamento do conhecimento da área e permitiu avaliar a pertinência das informações para a análise das características individuais e comuns de cada unidade terminológica.

5 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Esta seção apresenta a análise e discussão dos resultados, descritas a partir de exemplos resumidos - excertos da coleta de dados - em forma de quadros seguidos de comentários. São apresentadas as comparações que estabelecem as diferenças e semelhanças entre os termos estudados nos seus contextos de uso, evidenciando a presença de relações conceituais, imprescindíveis à estruturação de um sistema nocional.

Numa primeira leitura dos trechos identificados com o recurso “localizar e marcar” observou-se que o termo “*performance*” teve um número de ocorrências bem maior que o termo “*performance* musical”. O quadro abaixo demonstra os resultados da coleta, sendo que o número de ocorrências do termo *performance* musical está incluído no número de ocorrências do termo *performance*, uma vez que o recurso utilizado para localizar os termos não fez a distinção entre a palavra *performance* e a expressão *performance* musical.

QUADRO 1 - Número de ocorrências dos termos

FONTE	PERFORMANCE + PERFORMANCE MUSICAL	PERFORMANCE MUSICAL
ANAIS ANPPOM	903	69
MUSICA HODIE	97	22
OPUS	18	6
PERMUSI	430	48
TOTAL	1448	145

FONTE: Dados de pesquisa. Elaborado pela pesquisadora.

Desta forma, com o objetivo de ampliar o arcabouço de elementos sobre os termos, foram coletados os traços semânticos e características de uso dos termos: *performance* e *performance* musical. A análise que se segue baseia-se no estudo dos dois termos acima e apresenta as seguintes considerações:

A apuração do número de ocorrências, obtida pela análise de frequência, demonstrou uma diferença expressiva entre o número de vezes em que o termo foi empregado com traços semânticos e contextos de uso significativos e o número de ocorrências vazias de significado, descontextualizadas e evidenciando a falta de traços semânticos nos termos encontrados em títulos e subtítulos das seções dos artigos, citações, referências, notas, e, também, no corpo do texto, porém, sem elementos expressivos para a sua caracterização.

Na análise dos trechos transcritos, constatou-se maior ocorrência de contextos explicativos em detrimento dos contextos definitórios e associativos. Este fato dificultou, em certa medida, a elaboração da definição dos termos estudados, visto que as informações levantadas não apresentaram características referenciais dos termos, no entanto, está pertinente com a literatura terminológica que, de acordo com Krieger (2004), prescreve vários tipos de definições, dentre elas, a definição por implicação, que acontece quando a palavra é utilizada num contexto explicativo. O QUADRO 2 apresenta excertos da ocorrência desses contextos.

QUADRO 2 – Excertos da ocorrência de variação terminológica

TERMO	VARIAÇÃO	CONTEXTO
Performance musical	Práticas de performance	[...] uma vez que ela reflete diretamente as práticas de performance
		[...] práticas de performance específicas do saxofone
		[...] práticas de performance do baião e do jazz, especialmente do beebop

Prática da interpretação musical	No Choro, assim como em qualquer outro tipo de manifestação de música popular, o estudo da prática da interpretação musical (performance) torna-se um desafio para trabalhos de natureza acadêmica
Práticas interpretativas	Buscando aproximar o estudo científico de fontes musicais da prática interpretativa, as edições das obras do maestro Heinz Geyer receberam um texto inicial, com base em pesquisas históricas de jornais da época, programas de concertos e textos anteriormente publicados
Prática instrumental	[...] é o produto final da prática instrumental. No centro temos a “Performance”, que representa o produto final do processo – entendido aqui como o momento de apresentação pública

FONTE: Dados de pesquisa. Elaborado pela pesquisadora.

O emprego de diferentes unidades terminológicas com o significado de *performance* musical evidencia a presença da relação de equivalência entre os termos, indicando que o termo estudado é, também, representado por outros termos, devido às diferentes falas dos especialistas.

Estas divergências na comunicação se ancoram nos fundamentos da TCT (Cabré, 1993; 1999), que reconhece a existência da variação conceitual e denominativa nos domínios de especialidade e leva em conta a dimensão textual e discursiva dos termos.

Sendo assim, torna-se necessário, no controle de vocabulário, explicitar estas diferenças para contemplar as diversas possibilidades de falas dos especialistas e reconhecer a existência da variação terminológica, para se garantir uma comunicação de melhor qualidade. O QUADRO seguinte apresenta excertos da ocorrência de termos relacionados.

QUADRO 3 – Excertos da ocorrência de termos relacionados

TERMO	CONTEXTO
Performance studies	Constitui atualmente um ramo de conhecimento que se materializou em número considerável de pesquisas e de publicações
Etnografia da performance	O conhecimento do contexto permite que as análises dos parâmetros musicais sejam mais facilmente realizadas e compreendidas, porque abordadas a partir do conhecimento do ambiente musical do Choro, composto não somente pela música como também por inúmeros outros elementos
Espaço de performance	[...] em cada região do espaço de performance. Podemos notar certa homogeneidade, tanto quantitativa (número de instrumentos) quanto qualitativa (timbre e registro) da distribuição orquestral
	Outra importante abordagem de tratamento espacial na obra refere-se à distribuição de padrões em diferentes localidades do espaço de performance

Performance audiovisual	Por performance, compreende-se o resultado do conjunto significativo do gênero musical que se torna presença no videoclipe
	Os videoclipes são curtas metragens de técnica midiática que tecem formas de narrativas como música, poema, imagens, movimentos em diferentes contextos de performance audiovisual
Performance em grupo	No caso de conjuntos instrumentais, a coordenação dessas manipulações entre os intérpretes é crucial, tanto para a realização da performance quanto para o bom entendimento da ideia musical a ser transmitida. Músicos precisam prever as variações temporais produzidas pelos outros membros do conjunto
	Pode-se argumentar ainda que, no caso de performances em grupo, as informações visuais possuem papel essencial na sincronização dos eventos musicais. No entanto, músicos também são capazes de acompanhar outros músicos exclusivamente a partir de sinais acústicos, uma situação de performance que se torna cada vez mais comum como consequência do crescimento da prática de gravações de performances em estúdio, nas quais frequentemente um músico precisa seguir os outros através de fones de ouvido

FONTE: Dados de pesquisa. Elaborado pela pesquisadora.

A análise dos contextos de usos descritos acima, demonstra o emprego de diversas expressões que possuem características comuns com o termo *performance* musical, evidenciando as relações semânticas existentes entre os termos estudados e os termos relacionados, ampliando significativamente seu campo definicional.

A literatura (TÁLAMO; LARA; KOBASHI, 1992) mostra tratar-se de um tipo de relação associativa, que acontece quando uma conexão temática entre os termos pode ser estabelecida em virtude das características comuns compartilhadas entre os termos.

Uma vez que os termos relacionados remetem para outras possibilidades de significação associadas ao termo *performance* musical, é desejável que, no controle de vocabulário, estes termos sejam explicitados, a fim de se garantir aprofundamento no processo de apreensão do conhecimento especializado. O próximo quadro apresenta excertos da ocorrência de ambiguidades dos termos encontradas no mesmo artigo.

QUADRO 4 – Excertos da ocorrência de ambiguidades no mesmo artigo - Exemplo 1

TERMOS	CONTEXTO	FONTE
Performance musical	[...] é o produto final da prática instrumental	PM-2012-26-09
	[...] ato momentâneo da apresentação musical	
	[...] execução como segunda etapa de estudo	

FONTE: Dados de pesquisa. Elaborado pela pesquisadora.

Nestas ocorrências, as designações atribuídas ao termo pelo especialista, no mesmo artigo, evidenciaram uma relação de contiguidade no tempo, uma vez que foram empregadas para representar momentos distintos da *performance* musical. No trabalho terminográfico, a

análise dos aspectos das designações pode contribuir para a elaboração de uma definição mais completa do termo. A seguir, o quadro cinco apresenta outro exemplo de ambiguidades encontrados no mesmo artigo.

QUADRO 5 – Excertos da ocorrência de ambiguidades no mesmo artigo – Exemplo 2

TERMOS	CONTEXTO	FONTE
Performance musical	[...] processo de realização sonora onde o músico comunica aos ouvintes sua mensagem, interpretando determinado texto ou idéia musical	PM-2012-26-10
	[...] processo de mensuração do desempenho técnico do performer	

FONTE: Dados de pesquisa. Elaborado pela pesquisadora.

No exemplo acima, o especialista empregou o termo *performance* musical para definir dois processos distintos: interpretação musical e o processo de mensuração de desempenho técnico. Neste caso, o grupo de características levantadas determina o pertencimento do termo a categorias distintas do sistema nocional, evidenciando uma relação de associação, que permite ampliar o campo definidor do termo. De acordo com Lara (2001) não existe melhor ou pior definição para o termo, e sim, uma definição mais adequada à sua posição em relação ao sistema de conceitos. O próximo QUADRO apresenta ocorrências de ambiguidades dos termos em artigos diferentes.

QUADRO 6 – Excertos da ocorrência de ambiguidades em artigos diferentes

TERMOS	CONTEXTO	FONTE
Performance musical	[...] performance como reprodução de um texto	MH-12-2-13
	[...] atinente à experiência viva, ao <i>hit et nunc</i> do palco, à gestualidade e à aspectos corporais do músico-intérprete com relação aos modos e aos meios de sua apresentação com o instrumento	PM-2012-26-01
	[...] fenômeno que envolve não apenas o comportamento do instrumentista frente ao texto que interpreta, mas também os mecanismos de percepção envolvidos na escuta	OP-18-2-8

FONTE: Dados de pesquisa. Elaborado pela pesquisadora.

Os trechos transcritos de artigos publicados em diferentes periódicos da área, demonstraram a variedade de elementos que caracterizam o termo *performance* musical.

A presença de diversos aspectos nas designações do termo evidencia a adoção da linguagem natural pelos autores e sua análise pode contribuir para a identificação de

elementos nocionais que contribuem para uma definição mais elaborada. O próximo QUADRO também apresenta excertos da ocorrência de ambiguidades do termo *performance*.

QUADRO 7 – Excertos da ocorrência de ambiguidades do termo *performance*

TERMO	DOMÍNIO	CONTEXTO	FONTE
Performance	Artes	É um evento artístico e social	PM-2012-26-01
		[...] movimento e representação mímico-gestual do ator no palco	
		[...] remete à presença física no palco, ao corpo e a voz, ao meio e modo de interagir com o público espectador	
		[...] apresentação, execução, realização e funcionamento, as condições externas e internas da representação artística como um todo	
Performance	Esportes	[...] questões da performance, desenvolvendo exercícios para o detalhamento e aperfeiçoamento das cenas criadas	PM-2012-26-01

FONTE: Dados de pesquisa. Elaborado pela pesquisadora.

Verificou-se que o termo *performance* está caracterizado por elementos que o remetem para dois outros domínios do conhecimento: Artes e Esportes, evidenciando a sua polissemia, característica presente quando o termo é identificado por dois ou mais conjuntos de atributos.

Neste caso, para um controle eficaz do vocabulário, a literatura recomenda o uso de qualificadores que, segundo Gomes; Campos e Guimarães (2010) vão indicar a que domínio pertence o termo empregado eliminando, desta forma, os casos de ambiguidades. Observou-se que apesar do artigo ter sido publicado em periódico científico da área de Música, apresenta conceitos de outras áreas do conhecimento, como Artes e Esportes, o que evidencia a interdisciplinaridade do campo da Música.

Encerrando a etapa de análise dos trechos transcritos procedeu-se uma segunda etapa de análise da terminologia que teve como objetivo avaliar, através de amostragem, o emprego dos termos coletados na representação temática dos documentos. Desse modo, foram realizadas buscas em diferentes catálogos em linha de bibliotecas de instituições com programa de pós-graduação em música. Pretendeu-se verificar a contiguidade e semelhança entre o vocabulário utilizado na representação e o universo de conhecimento expresso nos textos.

O quadro seguinte mostra as variações terminológicas e termos relacionados obtidos na coleta, considerados como significativos, pois apresentaram, em suas designações, traços semânticos em relação estreita com os termos estudados.

QUADRO 8 – Resultado da pesquisa realizada em catálogos eletrônicos

TERMOS/ OCORRÊNCIA	UFMG	UFRGS	UFBA	UFRJ	UFPB	USP	UNESP	UNICAMP
Práticas de performance	0	0	0	0	0	0	0	0
Prática da interpretação musical	0	0	0	0	0	0	0	0
Práticas interpretativas	0	16	0	0	0	1	1	0
Prática instrumental	1	0	0	0	0	0	0	0
Performance studies	0	0	0	0	0	0	0	0
Etnografia da performance	0	0	0	0	0	0	0	0
Espaço de performance	5	0	0	0	0	0	0	0
Performance audiovisual	2	0	0	0	0	0	0	0
Performance em grupo	0	0	0	0	0	0	0	0

FONTE: Dados de pesquisa. Elaborado pela pesquisadora.

Foram pesquisados nove termos em catálogos eletrônicos de bibliotecas de instituições de ensino em música, que possuem doutorado em seu programa de pós-graduação. Delimitou-se a modalidade doutorado por considerar-se que neste nível de formação acadêmica espera-se encontrar maior especialização e aprofundamento do conhecimento. O levantamento foi realizado em julho de 2013, na modalidade de pesquisa avançada, no campo de assunto, na opção de termo “exatamente igual” ao termo pesquisado. São as seguintes as considerações sobre os resultados das buscas:

a) Os termos práticas de *performance*, prática da interpretação musical, etnografia da *performance*, *performance* em grupo e *performance studies*, não tiveram qualquer ocorrência nos catálogos consultados, evidenciando que os mesmos não são adotados como descritores nesses sistemas de informação.

b) O termo práticas interpretativas teve 16 ocorrências na UFRGS, 01 ocorrência na USP e 01 ocorrência na UNESP. Foi verificado que este é o termo autorizado no catálogo da UFRGS, que anula o termo *performance* musical, remetendo-o para o termo práticas interpretativas. Nos catálogos da USP e UNESP, os dois termos são adotados para

representar o mesmo documento, sem anular nenhum dos dois cabeçalhos. Esta ocorrência evidencia a relação de equivalência analisada no quadro dois.

c) O termo prática instrumental teve uma ocorrência no catálogo da UFMG, onde o documento também foi representado pelo termo *performance* musical, repetindo a mesma situação do item anterior.

d) O termo espaço de *performance* obteve cinco ocorrências no catálogo da UFMG. Analisando os registros bibliográficos, verificou-se que os documentos indexados não estão representados pelo termo “espaço de *performance*” e sim pelos termos *performance* (arte) e espaço (arte) evidenciando que existe uma interseção semântica entre eles, que poderia ser mais bem trabalhada num sistema de relações conceituais.

e) O termo *performance* audiovisual obteve duas ocorrências no catálogo da UFMG e demonstrou a mesma situação do item anterior. Os documentos foram representados pelos descritores arte audiovisual e *performance* (arte).

6 CONCLUSÃO

Esta pesquisa possibilitou demonstrar a importância de se realizar um estudo interdisciplinar da Ciência da Informação com a Terminologia, tanto no plano teórico como no plano prático. Constatou-se que a aproximação das duas ciências acontece por uma série de semelhanças observadas e que ambas perseguem o mesmo objetivo: resolver os problemas da comunicação do conhecimento no que refere ao tratamento dos dados informacionais.

A Ciência da Informação, área que visa organizar para transferir informação conta, para esta finalidade, com as linguagens documentárias, consideradas tradicionalmente como instrumentos de controle terminológico e que são adotadas em momentos distintos do processo de comunicação da informação. No primeiro momento - da representação da informação - o profissional da informação, mediante a terminologia da área, analisa, sintetiza e descreve o conteúdo dos documentos de uma determinada coleção, a fim de criar pontos de acesso para o referido documento, o que se constitui na atividade de indexação.

No segundo momento, também de acordo com a terminologia da área, o usuário formula suas estratégias de busca, a fim de acessar os documentos que tratam da informação desejada. Pode-se dizer, então, que a terminologia constitui a base dos vocabulários controlados, visto que a expressão dos conceitos se efetua através de termos.

Considerando que as linguagens documentárias são instrumentos constituídos por termos ou termos-descritores como visto na literatura (Tálamo; Lara; Kobashi, 1992) e que são instrumentos de controle terminológico para atuar em domínios específicos que fazem uso

das linguagens de especialidade, torna-se imprescindível a apropriação dos aportes da Terminologia para fundamentar teoricamente a criação desses instrumentos.

A Terminologia, área que visa estudar as linguagens de especialidade veiculadas nos documentos técnico-científicos conta com procedimentos metodológicos para a compreensão de um domínio, através do recolhimento e organização dos termos e o estabelecimento de relações semânticas para a organização do campo nocional. Quando um conjunto terminológico recebe um tratamento terminográfico, a questão das relações de significação assume grande importância, uma vez que a definição das unidades terminológicas é feita a partir dos discursos e o termo adquire um conceito específico e relacional, possibilitando, desta forma, a precisa identificação dos subdomínios tratados.

O trabalho desenvolvido nesta pesquisa teve a Música como área de aplicação dos fundamentos da Teoria Comunicativa da Terminologia (TCT), especificamente no subdomínio da *Performance* Musical. É oportuno salientar que o levantamento de toda a terminologia dessa área de especialidade não foi o objetivo deste trabalho, que, a seguir, relaciona algumas considerações que podem servir de norte na busca de dados mais aprofundados em trabalhos futuros, e que merecem especial atenção.

Uma das maiores dificuldades deste trabalho foi em relação à obtenção, no corpus selecionado, de contextos definitórios dos termos. Inicialmente, na análise do *corpus*, pressupôs-se que a leitura dos resumos e palavras-chave dos artigos seria suficiente para identificar a terminologia do subdomínio da *Performance* Musical. Não obstante, como resultado, obteve-se uma lista de termos vazios de significados, que não contribuíram para o estudo do campo nocional em questão, visto que a grande maioria das palavras-chave não constava do corpo do texto e, quando constava, não apresentava elementos significativos para compor a sua definição.

É necessário considerar que a distinção entre a linguagem natural e a linguagem de especialidade não se deu de maneira muito clara, visto que a linguagem de especialidade nada mais é do que um subconjunto do sistema linguístico – portanto, parte da língua – usado numa área de assunto particular e com uma terminologia específica. A presença de elementos característicos dos termos é imprescindível para a categorização dos mesmos.

A escolha do termo *performance* musical justificou-se pelo número significativo de sua ocorrência no *corpus* selecionado e pela necessidade de se obter elementos consubstanciados para a definição e delimitação do campo de estudo. A partir da análise dos contextos de uso com ocorrência dos termos *performance* e *performance* musical, emergiram outras unidades terminológicas, cujos traços semânticos evidenciaram a presença de um

vocabulário variado e estruturado com elementos nocionais, que contribuíram de forma significativa para a compreensão da área estudada.

O trabalho não se limitou a repertoriar apenas os termos selecionados, através do estudo de um termo específico, identificou-se um conjunto de termos significativos, extraídos de um *corpus* discursivo que uma vez categorizados poderão constituir um conjunto terminológico homogêneo a ser trabalhado como uma hipótese de organização do vocabulário controlado da área.

Foram registrados termos sinônimos, variações terminológicas e termos relacionados, considerados importantes para o entendimento das relações conceituais estabelecidas entre os termos e suas zonas de interseção semântica, ampliando, de forma expressiva, as possibilidades de significação no campo conceitual. A coleta manual dos trechos de ocorrência dos termos possibilitou à pesquisadora a leitura e subsequente apreensão de conceitos até então submersos na literatura especializada, resultando num aprofundamento do conhecimento da área.

Por fim, considera-se que o trabalho terminológico cumpre com o objetivo de contribuir para a definição de campos nocionais de áreas do conhecimento que não contam, ainda, com obras referenciais para a elaboração das linguagens documentárias. No momento em que são contempladas todas as manifestações emergentes nos discursos, às quais estão intimamente relacionadas aos termos delimitados *a priori* para a realização do estudo, o trabalho proporcionou uma navegação pelo sistema conceitual cognitivamente orientada.

Conclui-se que a adoção dos fundamentos da TCT pode servir como apoio à construção das linguagens documentárias porque, ao organizar os termos a partir das linguagens de especialidade, opera-se com as palavras em funcionamento, à medida que provê fundamentos para a estruturação e organização dos termos. Desta forma, pode-se garantir não só a comunicação precisa do conhecimento à comunidade usuária, como também tornar efetiva a transferência da informação e a geração de novos conhecimentos.

Espera-se que a presente pesquisa contribua para melhor organização, representação e recuperação da informação musical, sobretudo na harmonização das terminologias da área e que sirva de parâmetro para o desenvolvimento de novos trabalhos terminológicos em diferentes áreas do conhecimento.

REFERÊNCIAS

ASSIS, A.C. *et al.* Música e história: desafios da prática interdisciplinar. In: BUDAZ, R. (Org.). **Pesquisa e música no Brasil: métodos, domínios, perspectivas.** Goiânia: ANPPOM,

2009. Disponível em: <http://www.anppom.com.br/editora/Pesquisa_em_Musica-01.pdf>. Acesso em: 21 abr. 2013.

BÉLANGER, G. Pour une nouvelle terminographie. **Meta**, v.36, n.1, p.49-54, 1991.

BENVENISTE, E. Gênese do termo “cientifique”. In: _____. **Problemas de linguística geral, II**. Campinas: Pontes, 1989, p. 252-258.

BORÉM, F.; RAY, S. Pesquisa em performance musical no Brasil no século XXI: problemas, tendências e alternativas. In: SIMPÓSIO BRASILEIRO DE PÓS-GRADUANDOS EM MÚSICA, 2., 2012, Rio de Janeiro. **Anais...** Rio de Janeiro: UNIRIO, 2012.

CABRÉ, M. T. **La terminología**: representación y comunicación: elementos para una teoría de base comunicativa y otros artículos. Barcelona: Institut Universitari de Linguística Aplicada, 1999.

CABRÉ, M. T. **La terminología**: teoría, metodología, aplicaciones. Barcelona: Antártida/Empúries, 1993.

CINTRA, A.M.M. *et al.* **Para entender as linguagens documentárias**. 2. ed. rev. ampl. São Paulo: Polis, 2002.

GALVÃO, M.C.B. A linguagem de especialidade e o texto técnico-científico: notas conceituais. **Transinformação**, v. 16, n. 3, p. 241-251, set./dez. 2004.

GARDIN, J.C. **La logique du plausible**: essais d'épistemologie pratique. Paris : Maison des Sciences de l'Homme, 1987.

KOBASHI, N.Y. Resumos documentários: uma proposta metodológica. **Revista de Biblioteconomia de Brasília**, v. 21, n. 2, jul./dez. 1997.

KRIEGER, M. G.; FINATTO, M. J. B. **Introdução à terminologia**: teoria e prática. São Paulo: Contexto, 2004.

LARA, M. L. G. **A representação documentária**: em jogo a significação. 1993. Dissertação (Mestrado em Ciências da Comunicação) – Universidade de São Paulo, São Paulo, 1993.

LARA, M. L. G. Linguagem documentária e terminologia. **Transinformação**, v. 16, n. 3, p. 231-240, set./dez. 2004.

LARA, M. L. G. O processo de construção da informação documentária e o processo de conhecimento. **Perspectivas em Ciência da Informação**, v. 7, n. 2, jul./dez. 2002.

NOVELLINO, M.S.F. Instrumentos e metodologias de representação da informação. **Informação & Informação**, v. 1, n. 2, jul./dez. 1996.

PAVEL, S.; NOLET, D. **Manual de terminologia** [on line]. Canadá, 2002. Disponível em: <<http://www.translationbureau.gc.ca/>>. Acesso em: 14 fev. 2013.

ANÁLISE DE ONTOLOGIAS PARA REPRESENTAÇÃO DE DOCUMENTOS MULTIMÍDIA NA WEB

ANALYSIS OF ONTOLOGIES FOR REPRESENTING MULTIMEDIA DOCUMENTS ON THE WEB

Daniela Lucas Silva
Renato Rocha Souza

Resumo: O artigo aborda a situação de crescimento significativo de dados semanticamente relacionados e distribuídos na Web, em que padrões de metadados recomendados pelo *World Wide Web Consortium* (W3C) vêm sendo utilizados para descrever e representar recursos multimídia, possibilitando ampliar os pontos de acesso e melhorar a gestão, a organização e a recuperação de acervos digitais. O presente artigo buscou avançar nas pesquisas sobre representação semântica de documentos não necessariamente textuais, mas envolvendo também vídeos, imagens, áudios, propondo uma análise comparativa entre ontologias nessa perspectiva, as quais buscam organizar sistematicamente tipologias de metadados existentes para descrição de conteúdo multimídia frente a variados contextos e necessidades. A proposta da estrutura comparativa pode servir de instrumento facilitador a uma avaliação criteriosa em dimensões concernentes a reuso de ontologias, o que pode assegurar a seleção de recursos de conhecimento adequados para possíveis alinhamentos endereçados ao desenvolvimento de conceituações ou modelagens conceituais envolvendo ontologias para a Web.

Palavras-chave: Documentos multimídia. Padrões de metadados. Ontologias. Ontologias multimídia.

Abstract: The article discusses the situation of significant growth of semantically related data and distributed Web, where metadata standards recommended by the World Wide Web Consortium (W3C) have been used to describe and represent multimedia features, enabling the access points extend and improve management, organization and retrieval of digital archives. This paper aims to advance research on semantic representation does not necessarily textual documents, but also involving videos, images, audio, proposing a comparative analysis between ontologies that perspective, which seek to systematically organize typologies of existing metadata for describing multimedia content across the varying contexts and needs. The proposal of a comparative framework can serve as a facilitator to a careful evaluation regarding the reuse of ontologies dimensions, which can ensure the selection of adequate knowledge resources for possible alignments addressed the development of conceptual modeling involving ontologies for the Web.

Keywords: Multimedia documents. Metadata standards. Ontologies. Multimedia Ontologies.

1 INTRODUÇÃO

O crescimento exponencial de informações, ocasionado principalmente pelas facilidades introduzidas pelas tecnologias da informação e comunicação, vem impondo desafios no processo de produção, organização e disseminação de informação em unidades de informação como arquivos, bibliotecas, museus, centros de documentação e projetos de memória. Tais desafios encontram-se principalmente na descaracterização do formato de produção e consumo de informação associado a seus usuários. Na visão de Dodebei (2009), a

descharacterização é decorrente da digitalização das coleções, pois sua disponibilização na Internet rompe com o controle do equilíbrio entre oferta e demanda de informações e transfere para o ambiente digital o problema da gestão de informação. Segundo a autora, a sociedade se depara com vários projetos que buscam garantir o acesso a tudo que já foi produzido pela humanidade e há uma mobilização de recursos financeiros, humanos e materiais para garantir o direito à memória.

Uma das principais mudanças que reflete a Web é a possibilidade de integração e acesso a documentos digitais em diferentes suportes (texto, imagem, som, vídeo) por meio da inserção das tecnologias hipertexto e hiperímídia. Um fator de inovação que reflete as mídias digitais são os dispositivos eletrônicos (*tablets*, câmeras digitais, filmadoras, telefones inteligentes, MP3 *players*, televisores) integrados à Web que permitem o consumo e a gestão de conteúdo digital multimídia, propiciando um crescimento da disponibilidade da mídia. Tal crescimento torna-se incontrolável sem o apoio de instrumentos de representação de informação para prover armazenamento, organização e recuperação de modo eficaz.

Metadados é a forma mais comumente empregada para agregar semântica a informações (GILLILAND-SWETLAND, 2000) com o propósito de facilitar a busca de recursos de informação. O uso de metadados é uma prática antiga na área de biblioteconomia (BODOFF, HUNG e BEN-MENACHEM, 2005; ABBAS, 2010) no processo de catalogação e indexação de recursos bibliográficos, especialmente com a adoção de padrões como o MARC e o Dublin Core (ISQ, 2010), além de modelos conceituais para auxiliar a organização de metadados dessa natureza como o FRBR (IFLA, 1998). Entretanto, a necessidade de inclusão de metadados para além de recursos bibliográficos, isto é, metadados destinados à descrição de conteúdo multimídia, especialmente na Web, viabilizou o esforço conjunto entre comunidades e entidades normativas interessadas em fornecer um *framework* comum de metadados para aplicações de mídias inteligentes. São os casos *do World Wide Web Consortium (W3C)* e *da International Organization for Standardization/ International Electrotechnical Commission (ISO/IEC)* que buscam soluções inteligentes para descrição de conteúdo multimídia processável por máquina e baseada em semântica por meio da Web Semântica (BERNERS-LEE, HENDLER e LASSILA 2001) e da Interface de Descrição de Conteúdo Multimídia (MARTÍNEZ, 2004), conhecida como MPEG-7¹⁵⁶. Contudo, apesar de a visão sobre metadados ser similar tanto para a comunidade MPEG quanto para a

¹⁵⁶ <http://mpeg.chiariglione.org/standards/mpeg-7>

comunidade W3C, algumas diferenças em relação a propósito são evidentes. A primeira foca uma interface comum para descrever material multimídia refletindo informação sobre o conteúdo, o que difere da segunda que foca na descrição subjacente ao conteúdo (ou na descrição do conhecimento sobre o recurso) por meio de tecnologias semânticas, especialmente ontologias.

As comunidades de biblioteca digital, representação do conhecimento e inteligência artificial (uma parte da comunidade interessada em raciocínio automatizado sobre mídias) que interpretam, manipulam e geram arquivos de mídia audiovisuais têm participado intensamente em projetos de pesquisas voltados a modelos e tecnologias para processamento de metadados (OSSENBRUGGEN, NACK e HARDMAN, 2004; NACK, OSSENBRUGGEN e HARDMAN, 2005; BODOFF, HUNG e BEN-MENACHEM, 2005). Tais pesquisas são endereçadas a problemas significativos envolvendo metadados textuais e no contexto multimídia, principalmente pela convergência de processos informacionais na Web, a saber: i) *custo*, cuja produção de metadados de alta qualidade é uma atividade dispendiosa e consome tempo; ii) *subjetividade*, em que anotadores geralmente têm visões específicas acerca do conteúdo e do contexto em que é usado; iii) *restritividade*, em que um esquema com pouca restrição (como campo em texto livre) proporciona, geralmente, uma terminologia subjetiva, inconsistente e pouco aproveitada pela máquina; iv) *longevidade*, cuja construção de um esquema de anotação para fins específicos dentro de um domínio e suficientemente genérico entre domínios é uma tarefa difícil; v) *privacidade*, cujos metadados podem prover informações privadas ou sigilosas, o que requer cuidado especial; e vi) *padronização*, em que há necessidade de padronização tanto no nível de sintaxe quanto no nível semântico; o primeiro caso garante que uma ferramenta consiga analisar formatos de outras; e o segundo certifica que por meio de inferências automáticas, ferramentas podem descobrir a quais conceitos compartilhados um grupo de termos se refere.

A comunidade de biblioteca digital usa metadados como auxílio na catalogação e na recuperação da informação em grandes coleções de documentos. O padrão Dublin Core é o comumente usado na comunidade com os seus 15 elementos de metadados e qualificadores específicos destinados principalmente a catálogos em bibliotecas, incluindo título, assunto e dados de criação da obra. Para o W3C, a Web Semântica é uma tentativa de produzir resultados de pesquisas em biblioteca digital e representação do conhecimento que sejam aplicáveis à Web.

Pesquisadores de representação do conhecimento ocupam-se com a infraestrutura tecnológica bem projetada para adicionar metadados a documentos Web com vistas à

publicação formal e explícita de conhecimento na Web. Ontologias são estruturas-chave para compartilhar e comunicar conhecimento explícito por meio de um compromisso ontológico entre pessoas e aplicações (GRUBER, 1993; GUARINO, 1998) que irão usar a terminologia da ontologia de acordo com regras especificadas. No que tange à Web, ontologias são concebidas por meio de linguagens de representação (ou declarativas) como OWL na especificação de objetos, conceitos, relacionamento entre eles e regras que limitam a interpretação destes em certo domínio. Como exemplo, cita-se um cenário (BODOFF, HUNG e BEN-MENACHEM, 2005) em que um vídeo na Web é descrito por metadados que especificam que o vídeo é sobre “dois jovens profissionais urbanos”; uma consulta feita com a sentença “estilo de vida yuppie” não retornaria a página. Combinando o metadado com uma ontologia, poder-se-ia declarar que “Yuppie” é um acrônimo comum para “Jovem Profissional Urbano” e que denota um estilo de vida específico. Desse modo, a mesma consulta retornaria a página com o vídeo em questão.

A norma ISO MPEG-7 empreendeu esforços na proposição de uma interface comum para descrever material multimídia, refletindo informação sobre o conteúdo. Entretanto, o MPEG-7 apesar de ser um padrão de descrição recomendado pela comunidade multimídia, principalmente por seu *background* de conhecimento neste domínio, possui limitações de ordem semântica (OSSENBRUGGEN, NACK e HARDMAN, 2004; NACK, OSSENBRUGGEN e HARDMAN, 2005; ARNDT et al., 2009; NIXON et al., 2011) por ser baseado no formato XML *Schema*. Tal característica pode acarretar um desacordo entre a informação que pode ser extraída do dado e a interpretação que o usuário faz numa dada situação envolvendo o mesmo dado, conhecida como *gap* semântico (SUÁREZ-FIGUEROA, ATEMEZING e CORCHO, 2013). Por outro lado, O W3C e comunidades afins vêm empregando esforços de pesquisas (ARNDT et al., 2009; STEGMAIER et al., 2009; NIXON et al., 2011; SUÁREZ-FIGUEROA, ATEMEZING e CORCHO, 2013) para ir além dos padrões de metadados correntes com a adoção de ontologias multimídia baseadas em RDF/OWL e no padrão MPEG-7.

Apesar das diferenças entre as abordagens sobre modelagem de informação nas comunidades supracitadas, muitas aplicações necessitam de elementos de pesquisa das áreas envolvidas. Segundo Ossenbruggen, Nack e Hardman (2004), ontologias podem ser endereçadas, por exemplo, para organização e controle da terminologia usada em metadados; além de auxiliar usuários na produção de anotações mais sistemáticas e consistentes por meio de conhecimento explícito acerca de um domínio. Contudo, surgem desafios na identificação e seleção de uma variedade de padrões de metadados, vocabulários e ontologias disponíveis e

que precisam ser compatíveis com as entidades reais de um domínio específico. Tais desafios encontram-se especialmente i) no alinhamento de vocabulários e ontologias que reflete aspectos de interoperabilidade semântica e sintática para o provimento de compartilhamento entre sistemas e aplicações na web; e ii) na modelagem conceitual adequada para representar consensualmente parte da realidade de um domínio.

Diante ao exposto, a proposta deste artigo é apresentar o resultado da análise comparativa envolvendo vocabulários para representação semântica de documentos não necessariamente textuais, mas envolvendo também vídeos, imagens, áudios, propondo o uso de ontologias para viabilizar o tratamento semântico de descritores para documentos multimídia em um dado domínio. A principal contribuição deste trabalho está em trazer à luz os padrões existentes (incluindo metadados, vocabulários e ontologias) endereçados à descrição de documentos de tipo multimídia para pesquisadores das ciências da informação e da computação. Por fim, a estrutura comparativa de propostas e padrões busca desvelar características que poderiam, deveriam e não estão sendo consideradas para caracterização desse tipo de recurso, principalmente no contexto da Web, refletindo nas decisões de seleção para reuso de recursos de conhecimento disponíveis.

O artigo está organizado como se segue. Nesta introdução, apresenta-se a motivação da pesquisa, contextualizando a situação problemática e delineando os objetivos a serem alcançados. O capítulo 2 descreve a metodologia de pesquisa, elucidando os métodos e as técnicas utilizados para o alcance dos objetivos. O capítulo 3 traz os resultados da análise comparativa das ontologias multimídia selecionadas na pesquisa a partir de critérios pré-determinados. Finalmente, o capítulo 4 se incumbem de traçar as conclusões do artigo.

2 METODOLOGIA DE PESQUISA

Dado que o objetivo geral deste artigo é apresentar uma análise comparativa entre ontologias para o domínio de anotação de documentos multimídia, fez-se necessária a adoção de métodos e técnicas voltados à identificação, seleção e elaboração de critérios de análise e comparativos envolvendo os objetos empíricos de pesquisa. Para tal, foi necessária uma revisão na literatura da área de engenharia de ontologias a fim de se encontrar um guia metodológico atual, testado e validado em diferentes domínios e áreas. O procedimento permitiu identificar e selecionar a *NeOn Methodology* (SUÁREZ-FIGUEROA, GÓMEZ-PÉREZ e FERNÁNDEZ-LÓPEZ, 2012), advinda de *frameworks* metodológicos amplamente aceitos em áreas maduras como engenharia de software e engenharia do conhecimento (FERNANDEZ, GOMEZ-PEREZ e JURISTO, 1997).

A metodologia de engenharia de ontologias NeOn abrange nove cenários¹⁵⁷ que sugerem uma série de passos flexíveis para o desenvolvimento de ontologias para a Web, focando o reuso de recursos de conhecimento disponíveis. Dentre os cenários indicados no guia, destaca-se neste artigo a utilização de parte do Cenário 3 - *reuso de recursos ontológicos* – que compreende o reuso de possíveis recursos ontológicos existentes para a construção ou aprimoramento de uma rede de ontologias. Um recurso ontológico abrange, por exemplo, ontologias já definidas ou partes de ontologias disponíveis e úteis à resolução de problemas. As etapas metodológicas deste Cenário efetuadas na pesquisa foram: i) busca por ontologias: buscar por recursos ontológicos candidatos à análise documental e de conteúdo, preferencialmente em repositórios da Web Semântica; e ii) comparação de ontologias: estabelecer uma análise comparativa entre os recursos ontológicos selecionados em (i), levando-se em consideração critérios determinados pelo ontologista, tais como clareza no código, qualidade no conteúdo, custo-benefício para reuso, etc.

O processo metodológico, assim como a descrição de seus métodos e técnicas, é elucidado nas seções subsequentes da seguinte forma: a seção 2.1 apresenta o método usado para identificação e seleção de ontologias multimídia na literatura e em repositórios específicos; e a seção 2.2 elucida a estratégia metodológica determinada para apoiar a análise e a comparação das ontologias selecionadas.

2.1 Identificação e seleção de ontologias multimídia

O primeiro passo foi realizar a atividade de aquisição de conhecimento sobre o assunto “descrição documental de conteúdo multimídia”, a qual envolveu um estudo acerca do domínio por meio de fontes documentais, incluindo normas, artigos e bibliotecas de esquemas XML relacionadas a padrões para descrição de documentos multimídia. Os padrões ISO MPEG-7 e Dublin Core foram selecionados como material de referência para aquisição de conhecimento sobre o domínio de descrição de recursos multimídia, pois, em revisão recente de literatura (SILVA e SOUZA, 2013), evidenciou-se que grande parte de ontologias multimídia é construída com fundamentos advindos do padrão MPEG-7 e do padrão Dublin Core. Para o levantamento de características multimídia foram consultados os descritores e os esquemas de descrição do MPEG-7¹⁵⁸ e os elementos do Dublin Core¹⁵⁹, além do documento

¹⁵⁷ Em virtude do limite de extensão do presente artigo, optou-se em não elucidar os nove cenários do guia. A descrição completa pode ser conferida em SUÁREZ-FIGUEROA, GÓMEZ-PÉREZ e FERNÁNDEZ-LÓPEZ (2012).

¹⁵⁸ `xmlns:mpeg7=urn:mpeg:mpeg7:schema:2001`

ISO/IEC JTC1/SC29/WG11N6828 (MARTÍNEZ, 2004) que especifica os requisitos do padrão MPEG-7.

A organização da aquisição de conhecimento ocorreu em três categorias de tipos de metadados fundamentados na literatura. Tais categorias buscam representar características sobre descrição de conteúdo multimídia e são descritas como se segue. A categoria *metadados independentes de conteúdo* é direcionada ao gerenciamento e a administração de recursos de informação e foi organizada em quatro tipos de descrição, a saber: i) criação e produção da mídia; ii) classificação da mídia; iii) informação da mídia; e iv) uso da mídia. A categoria *metadados dependentes de conteúdo* foi organizada nas subcategorias metadados visuais e metadados para áudio, ambas consideradas de baixo nível e geralmente seus conteúdos são extraídos automaticamente por algoritmos computacionais. E a categoria *metadados descritivos de conteúdo* se caracteriza por associar entidades da mídia com entidades do mundo real e foi organizada nas seguintes subcategorias: i) segmentos de mídia; ii) semântica de conteúdo; iii) personalização de conteúdo; e iv) características de alto nível envolvendo áudio.

O segundo passo, por sua vez, foi identificar ontologias multimídia fazendo-se um levantamento na literatura e buscas em repositórios da Web Semântica. O critério usado para a seleção de ontologias multimídia na literatura se baseou no conceito de “multimídia”, que reflete um documento composto que faz referência a vários tipos de objetos, tais como vídeo, texto, som, imagem, entre outros, e que pode ser dividido em partes que contêm textos, imagens, e assim por diante (SITARAM e DAN, 1999). Outro princípio adotado para o critério de seleção foi o de seguir diretrizes relevantes na literatura (SUÁREZ-FIGUEROA, ATEMEZING e CORCHO, 2013) e recomendadas pelo padrão de metadados MPEG-7 no aspecto de descrição multimídia, envolvendo: i) descrições no assunto expressando a semântica transmitida; ii) descrições estruturais permitindo a decomposição e localização das partes de conteúdo; e iii) descrições de baixo nível cobrindo características de áudio e visuais. Para a identificação e seleção de ontologias em repositórios Web, o guia NeOn recomenda o uso de máquinas de busca para a recuperação de ontologias em repositórios da Web Semântica. Nesse sentido, as máquinas de busca selecionadas, dentre as opções listadas pela comunidade *Linked Open Data*¹⁶⁰, foram o Watson¹⁶¹ e o Swoogle¹⁶² pelo fato de terem sido

¹⁵⁹ <http://dublincore.org/documents/dces/>

¹⁶⁰ <http://esw.w3.org/topic/TaskForces/CommunityProjects/LinkingOpenData/SemanticWebSearchEngines>

¹⁶¹ <http://watson.kmi.open.ac.uk/WatsonWUI/>

bem avaliadas em projetos e validações de uso (SUÁREZ-FIGUEROA, GÓMEZ-PÉREZ e FERNÁNDEZ-LÓPEZ, 2012). Os termos chave usados para a busca de documentos semânticos foram determinados a partir da atividade de aquisição de conhecimento anteriormente comentada. A análise de conteúdo das ontologias identificadas foi realizada por meio do próprio navegador Web, do editor Protégé 4.3 e, em casos específicos, do recurso OWLDoc¹⁶³, em que permite uma navegação taxonômica sobre as classes envolvidas.

As atividades de busca na literatura e em repositórios Web resultaram, num primeiro momento, em 17 ontologias multimídia candidatas a análise. Após um processo de refinamento¹⁶⁴ frente às ontologias previamente selecionadas para análise, uma lista unificada de 9 ontologias é apresentada no Quadro 1 em ordem cronológica de suas versões, informando o nome da ontologia, o *link* de acesso (com última versão e forma de acesso) e características de cobertura. Em seguida, uma descrição breve de cada ontologia é apresentada.

QUADRO 8 - Lista unificada: ontologias multimídia selecionadas para análise

Ref.	Ontologias multimídia	Links	Cobertura
1	Media Ontology	http://www.w3.org/TR/mediaont-10/ (versão 2013)	Áudio, vídeo, imagem.
2	M3 Multimedia	http://buscamedia.isoco.net/m3repository/M3OntologyNetwork/owldoc/index.html (versão 2012)	Imagem, vídeo, texto, áudio, multimídia.
3	M3O	http://m3o.semantic-multimedia.org/ (versão 2010)	Imagem, vídeo, texto, áudio, multimídia.
4	MCO e MDO Boemie	http://mklab.itι.gr/project/boemieontology (Versão 2008)	Áudio, vídeo, texto, imagem, audiovisual, multimídia.
5	COMM	http://multimedia.semanticweb.org/COMM/ (versão2007)	Áudio, vídeo, notícias, imagem, multimídia.
6	Polysema	http://polysema.di.uoa.gr/ont/mds.owl (versão 2007)	Áudio, audiovisual, imagem, multimídia, vídeo.
7	Hunter	http://metadata.net/mpeg7/mpeg7.owl (versão 2006; acesso por cache)	Áudio, audiovisual, imagem, multimídia, vídeo.
8	SmartWeb	http://www.smartweb-project.de/ontology/swinto0.3.1.rdfs (versão 2006)	Audiovisual

¹⁶² <http://swoogle.umbc.edu/>

¹⁶³ <http://buscamedia.isoco.net/m3repository/M3OntologyNetwork/owldoc/index.html>

¹⁶⁴ Tentativas de abertura das ontologias no editor Protégé, visualização de anomalias em suas estruturas, e ontologias fora do escopo proposto foram consideradas no sentido de justificar a exclusão de algumas delas da lista.

9	Rhizomik	http://rhizomik.net/ontologies/2005/03/Mpeg7-2001.owl (versão 2005)	Audiovisual
---	----------	---	-------------

Fonte: elaborado pela autora.

A *Media Ontology* foi proposta no ano de 2009 (STEGMAIER et al., 2009) dentro do *W3C Media Annotation Working Group*. A ontologia foi construída utilizando padrões de metodologias de engenharia de ontologias e seu propósito de construção foi de definir um conjunto de propriedades de anotação centrais para descrever conteúdo multimídia, juntamente com um conjunto de mapeamentos entre os principais formatos de metadados em uso atualmente.

A *M3 Multimedia* faz parte de uma ontologia abrangente denominada *M3 Ontology Network*¹⁶⁵, resultado de um ambicioso projeto de pesquisa espanhol denominado *Buscamedia*. O projeto objetiva criar um mecanismo de busca semântica de recursos multimídia visando progressos nas áreas de semântica, produção audiovisual e distribuição de mídia independente de redes de terminais e consumidores. No ano de 2012 foi finalizada a versão 2.7 da *M3 Ontology Network* como parte do projeto *Buscamedia*. A M3 tem como objetivo modelar informação multimídia para qualquer tipo de recurso em vários domínios e num contexto multi-idioma.

A *Multimedia Metadata Ontology* (M3O) (SAATHOFF e SCHERP, 2010) foi proposta em 2010 como um modelo abrangente para representar metadados multimídia, contemplando combinações de modelos e padrões de metadados para descrever semanticamente apresentações de documentos multimídia estruturados. O modelo M3O possui uma diferença em relação a outros projetos de ontologias multimídia em não focar na representação ontológica fundamentada em um padrão de metadado específico, como, por exemplo, o MPEG-7.

O projeto *Boemie* (*Bootstrapping Ontology Evolution with Multimedia Information*) foi organizado e constituído por membros vinculados a Centros de Pesquisa Europeus interessados em análise semântica multimídia. O projeto teve seu início no ano de 2006 e relatório final publicado no ano de 2008 (DASIOPOULOU et al., 2008). Almejou o desenvolvimento de ontologias multimídia e de domínios específicos com o propósito de representar semântica multimídia dentro de um cenário de aplicação integrada contemplando

¹⁶⁵ <http://buscamedia.isoco.net/m3repository/M3OntologyNetwork/M3OntologyNetwork.html>

representação e aquisição de conhecimento, raciocínio, evolução de ontologia multimídia, recuperação e apresentação.

A *Core Ontology for Multimedia*, conhecida como COMM, foi desenvolvida no ano de 2007 por um grupo de renomados pesquisadores nas áreas multimídia, bibliotecas digitais e Web Semântica (ARNDT et al., 2009). O propósito principal da ontologia COMM é fornecer uma conceituação fundamental (denominada *core ontology*) para descrição multimídia cobrindo de maneira genérica um domínio em específico que lida com conteúdo multimídia.

A ontologia MPEG-7 MDS (VALKANAS, TSETOS, HADJIEFTHYMIADES, 2007) foi desenvolvida no ano de 2006 dentro do projeto Polysema¹⁶⁶ conduzido por grupos de pesquisa da *University of Athens (Pervasive Computing Research Group)* e parcerias da indústria (ex. Siemens). O projeto propôs uma infraestrutura adequada para gerenciamento e processamento semântico de conteúdo multimídia com o uso de ontologias e padrões de metadados em ambientes interativos, especialmente serviços de TV digital e ferramentas para anotação de vídeos.

A ontologia MPEG-7 de Hunter foi proposta em 2001 dentro do projeto *Harmony International Digital Library*¹⁶⁷. Hunter (2001) propôs uma tradução manual do MPEG-7 em RDFS (então convertida em DAML-OIL) e posteriormente esta ontologia foi traduzida em OWL, estendida e harmonizada com a ontologia de alto nível ABC para endereçar descrições semânticas vinculadas a ontologias de domínios específicos.

O projeto SmartWeb¹⁶⁸ (VEMBU et al., 2006) foi conduzido entre os anos de 2004 e 2007 e combina conhecimento especializado de diversas áreas de conhecimento. Um conjunto de ontologias relevantes tem sido desenvolvido para apoiar sistemas móveis, multimodais e inteligentes com funcionalidades de respostas a consultas de usuários em temáticas de domínios diversos na Web.

Finalmente, a ontologia Rhizomik (GARCÍA e CELMA, 2005) foi desenvolvida dentro do projeto ReDeFer¹⁶⁹ em 2005 com o propósito de produzir ontologias fundamentadas no padrão MPEG-7 para integração com iniciativas de metadados multimídia existentes. O projeto teve uma abordagem diferenciada das propostas de tradução manual de partes do

¹⁶⁶ http://polysema.di.uoa.gr/index_en.html

¹⁶⁷ <http://metadata.net/harmony/index.html>

¹⁶⁸ http://smartweb.dfki.de/start_en.html

¹⁶⁹ <http://rhizomik.net/redefer>

padrão MPEG-7, como ocorreu na ontologia de Hunter. A ontologia Rhizomik propõe uma tradução automática completa do esquema MPEG-7 em OWL.

2.2 Análise e comparação de ontologias multimídia

A atividade de análise das ontologias multimídia foi considerada como um processo de auditoria, no qual as ontologias e documentações subjacentes foram inspecionadas e analisadas. Os critérios determinados para analisar as ontologias multimídia são, em sua maioria, oriundos do guia metodológico da NeOn, os quais foram originados de casos de uso do guia em diversas experiências de projeto envolvendo desenvolvimento e reúso de ontologias. Outros critérios foram concebidos e adaptados para esta pesquisa, que, em conjunto com os recomendados pelo guia, foram organizados em dimensões concernentes a reuso elucidadas como se segue: i) esforço para entendimento: estimativa de esforços necessários para entendimento do conteúdo da ontologia candidata; ii) esforço para integração: estimativa de esforços empreendidos para integrar a ontologia candidata à ontologia que está sendo construída; e iii) confiabilidade: análise da confiança em relação à ontologia candidata frente a aspectos de tratamento semântico nas declarações (ex. axiomas presentes; recursos de conhecimento utilizados), avaliação (ex. testes disponíveis) e projetos renomados que fazem uso.

Para a análise documental, os seguintes critérios foram direcionados: i) propósito de desenvolvimento da ontologia; ii) disponibilidade de documentação da ontologia; iii) recursos de conhecimento utilizados; iv) disponibilidade de conhecimento externo; v) disponibilidade de testes; vi) disponibilidade de resultados de avaliação de testes; vii) informação sobre a equipe de desenvolvimento; e viii) informação sobre projetos e ontologias que fazem uso. E para a análise de código das ontologias, os critérios empregados foram: i) representação de metadados multimídia; ii) linguagem de implementação; iii) clareza no código; iv) adequação a extração de conhecimento; v) adequação a convenção de nomes; e vi) anotações e axiomas existentes na terminologia. O QUADRO 2 apresenta e descreve os critérios em suas respectivas dimensões.

QUADRO 2 - Critérios para a análise de ontologias multimídia candidatas a reúso

Critérios	Descrição
Esforço para entendimento	
Qualidade da documentação	Relacionado à existência de qualquer material usado para descrever a ontologia candidata como, por exemplo, decisões de modelagem. A partir deste material, um não especialista do domínio deve entender o conhecimento representado na ontologia.
Disponibilidade de conhecimento externo	Relacionado à referência de fontes documentais externas em projetos que fazem uso da ontologia candidata e/ou acesso fácil a autores e especialistas.
Clareza no código	Relacionado à facilidade no entendimento e modificação do código; se as entidades contidas no

	código seguem um padrão, se são claras e coerentes; se existem comentários; e se o código é documentado.
Anotações identificadas na terminologia	Relacionado à existência e à qualidade das anotações realizadas nos elementos da terminologia da ontologia candidata, promovendo informações relevantes sobre os mesmos.
Esforço para integração	
Representação de metadados multimídia	Relacionado a características multimídia satisfeitas pelas ontologias candidatas a reuso, apoiando, assim, o processo de reuso de conceitos multimídia adequados a características provenientes dos padrões MPEG-7 e Dublin Core.
Adequação a extração de conhecimento	Relacionado à facilidade de identificação e extração de partes do conhecimento na ontologia candidata a reuso.
Adequação a convenção de nomes	Relacionado à verificação de regras associadas à determinação de nomes (terminologia) dos componentes ontológicos (conceitos, relações, etc.).
Adequação a linguagem de implementação	Relacionado à verificação da linguagem que representa os componentes da ontologia candidata, isto é, se se trata de uma linguagem compatível à almejada para representar conhecimento similar e com a mesma granularidade.
Confiabilidade	
Disponibilidade de testes	Relacionado à disponibilidade de testes para a ontologia candidata a reuso.
Avaliação de testes	Relacionado à existência de um conjunto de unidades de testes usado na avaliação da ontologia candidata.
Reputação do time de desenvolvimento	Relacionado à reputação da equipe de desenvolvimento da ontologia candidata.
Confiabilidade no propósito	Relacionado ao propósito de desenvolvimento da ontologia candidata. Por exemplo, ontologias desenvolvidas como uma amostra acadêmica simples são menos confiáveis do que àquelas desenvolvidas para serem usadas em projetos reais.
Suporte prático	Relacionado à existência de outros projetos conhecidos ou ontologias que reusam a ontologia candidata.
Recursos de conhecimento utilizados	Relacionado a recursos ontológicos (ex. ontologias de fundamentação, etc.) e não ontológicos (ex. padrões de metadados) usados na ontologia candidata.
Axiomas identificados na terminologia	Relacionado à existência de axiomas nos elementos da ontologia, garantindo, assim, restrições acerca de suas interpretações.

Fonte: elaborado pela autora.

3 RESULTADOS DA ANÁLISE COMPARATIVA DAS ONTOLOGIAS MULTIMÍDIA

A partir da análise comparativa realizada por meio de critérios metodológicos foi possível delinear considerações relevantes sobre as ontologias multimídia selecionadas para estudo bem como as dimensões concernentes a reuso: esforço para entendimento, esforço para integração e confiabilidade. As evidências são elucidadas a seguir.

A dimensão *esforço para entendimento* apresentou-se com avaliações baixas para algumas ontologias analisadas. Contribuíram para essa realidade a Polysema MPEG-7, a MPEG-7 Rhizomik, a MPEG-7 Hunter e a M3 Multimedia, em geral por motivos de escassez de fontes documentais e/ou ausência de anotações, ou mesmo sem contribuição semântica, nos elementos de suas estruturas. Tal constatação desqualifica essas ontologias no aspecto de um consumo considerável de tempo para se conseguir entender seus propósitos, escopos e conceituações visando alinhamentos consistentes.

A dimensão *esforço para integração* foi bem qualificada em comparação com as outras duas e também a que se manteve mais estável. Isso demonstra que as nove ontologias multimídia se apresentaram como boas candidatas a reuso em relação a esforço empreendido para alinhamentos. Reforça também que o método aplicado na seleção de ontologias

multimídia a compor o corpus da pesquisa foi bem sucedido. Torna-se válido ressaltar que um de seus critérios, representação de metadados multimídia, possibilitou algumas observações relevantes acerca da existência de lacunas que podem ser endereçadas à proposição de recomendações para modelagem multimídia. São elas: metadados para classificação e uso da mídia, personalização de conteúdo e características auditivas de alto e baixo nível são pouco explorados no contexto das ontologias analisadas. Descritores MPEG-7 específicos para áudio, tanto de alto nível quanto de baixo nível, são cobertos apenas pela ontologia MPEG-7 Rhizomiki. Algumas, como são os casos da MDO Boemie e SmartWeb MPEG-7, modelam classes de recursos de áudio desprovidas de axiomas e anotações para comportar futuras implementações.

A dimensão *confiabilidade* pode ser considerada uma característica presente na maioria das ontologias analisadas pelas seguintes constatações: i) todas possuem uma equipe de desenvolvimento com boa reputação; ii) todas são assistidas por entidades importantes no cenário mundial, tais como W3C, European Commission, German Federal Ministry of Education and Research, universidades europeias conceituadas e renomados centros de pesquisa; e iii) grande parte (M3O, COMM, Boemie, M3 Multimedia, Rhizomik) se propôs a disponibilizar ricas axiomatizações em suas conceituações, as quais são fundamentadas, na maioria dos casos, em ontologias de alto nível, em padrões de projeto multimídia, e no padrão MPEG-7.

Ontologias de fundamentação e padrões de projeto multimídia são recursos de conhecimento usados efetivamente apenas pela COMM e M3O na organização semântica de seus elementos multimídia. A M3 Multimedia também o faz de forma indireta quando estabelece reuso de toda a estrutura de conhecimento da COMM, porém não deixa claro em sua escassa documentação como usufrui dos padrões multimídia da COMM em sua conceituação. E a SmartWeb MPEG-7 apesar de usar ontologias de fundamentação para organização de entidades semânticas provenientes da integração com ontologias de domínio, não possui uma estrutura consistente para organizar semanticamente características multimídia como segmentação e anotação.

Estratégias diversificadas de modelagem para metadados descritivos de conteúdo foi um aspecto notório na análise das ontologias multimídia. Algumas ontologias propagam ambiguidades semânticas presentes no padrão MPEG-7 buscando seguir a flexibilidade de seus esquemas; outras buscam restringir classes e propriedades relacionadas ao padrão por meio de axiomas, como ocorreu com a MCO Boemie. Há de se pensar na estratégia mais conveniente numa perspectiva de uso dessas ontologias na Web em que as possibilidades

interpretativas dos agentes (humanos e computacionais) são diversas. E, por fim, ontologias para anotação semântica são concebidas como meio de a máquina interpretar os metadados com semânticas bem definidas e necessárias para garantir que o anotador e o consumidor da anotação compartilhem o mesmo significado perante um recurso multimídia.

Finalmente, os QUADROS 3 e 4 apresentam um resumo da análise comparativa de conteúdo documental das ontologias multimídia. E o QUADRO 5 apresenta um resumo da análise comparativa envolvendo a inspeção do código das ontologias multimídia.

QUADRO 3 - Quadro sinóptico da análise de conteúdo documental (a)

categorias de análise de conteúdo	Media Ontology	COMM	Boemie	M3 Multimedia	M3O
Propósito de desenvolvimento	Definir descritores centrais para descrever conteúdo multimídia.	Cobrir genericamente um domínio que lida com conteúdo multimídia.	Representar semântica multimídia dentro de um cenário de aplicação integrada.	Criar um mecanismo de busca semântica de recursos multimídia.	Integrar modelos e padrões de metadados por meio de um <i>framework</i> genérico de modelagem semântica.
Disponibilidade de documentação	Informação sobre as propriedades de anotação determinadas.	Informação sobre os padrões multimídia especializados dos padrões de projeto da DOLCE.	Informações sobre a modelagem das conceituações das ontologias multimídia.	Informações sobre o projeto Buscamedia e sua perspectiva multimídia.	Informações sobre a modelagem dos padrões multimídia e de proveniência.
Recursos de conhecimento utilizados	Padrões de metadados W3C para mídia, destacando os padrões MPEG-7 e Dublin Core.	Padrão MPEG-7; DOLCE e padrões de projeto D&S e OIO.	Padrão MPEG-7	DUL; LIR; Media Ontology; Boemie VDO; COMM; FOAF; Dublin Core.	DUL e padrões de projeto DnS, <i>Information and Realization e Data Value</i> .
Disponibilidade de conhecimento externo	Lista de casos de uso compilados envolvendo vocabulários multimídia, com destaque à disponibilidade das especificações de requisitos que fundamentam a ontologia.	Organização conceitual (OWLDoc) de cada módulo da COMM reusado na ontologia <i>M3 Multimedia</i> .	Organização conceitual (OWLDoc) da MDO reusada na ontologia <i>M3 Multimedia</i> .	Sítio do <i>Ontology Engineering Group</i> .	Sítio sobre o <i>framework</i> SemanticMM4U e sua integração com a M3O.
Disponibilidade de testes	Conjunto de testes envolvendo todos os formatos contemplados no documento de especificação da ontologia.	Ausente	Ausente	Ausente	Informação sobre o método de integração de ontologias e padrões de metadados em casos reais de aplicação.
Disponibilidade de resultados de avaliação de testes	Arquivos de compatibilização de vocabulários disponíveis, testados e avaliados.	Ausente	Ausente	Ausente	Informação sobre os resultados de avaliações sobre os testes de integração.
Equipe de desenvolvimento	<i>W3C Media Annotation Working Group</i>	<i>W3C Multimedia Annotation Interoperability</i> . Ontologia referenciada no <i>W3C Multimedia Vocabularies on the Semantic Web</i> .	Centros de Pesquisa Europeus	<i>Ontology Engineering Group</i>	Grupo de pesquisa acadêmico
Projetos e ontologias que fazem uso	Projeto Buscamedia e sua ontologia multimídia denominada <i>M3 multimedia</i> ; Multimedia Metadata Ontology (M3O).	Projetos X-Media, K-Space, Buscamedia e Organizações Globo. Multimedia Metadata Ontology (M3O).	Projeto Buscamedia e sua ontologia multimídia denominada <i>M3 multimedia</i> .	Projeto Buscamedia	Projeto WeKnowIt Projeto X-Media.

Fonte: elaborado pela autora.

QUADRO 4 - Quadro sinóptico da análise de conteúdo documental (b)

Categorias de análise de conteúdo	MPEG-7 Hunter	MPEG-7 Rhizomik	SmartWeb MPEG-7	Polysema MPEG-7 MDS
Propósito de desenvolvimento	Propor uma tradução manual do padrão de metadado MPEG-7 visando integração com tecnologias da Web Semântica.	Produzir ontologias por meio de tradução automática e completa do padrão MPEG-7 para integração com iniciativas de metadados multimídia existentes.	Produzir ontologias fundamentadas no padrão MPEG-7 para anotação de conteúdo multimídia envolvido em dispositivos inteligentes e móveis.	Cobrir descrições baseadas no MDS MPEG-7 e subsidiar a ferramenta de anotação de vídeo.
Disponibilidade de documentação	Informação sobre as classes e propriedades relacionadas a conteúdo multimídia, decomposição de segmentos e descritores visuais.	Informação sobre a metodologia empregada para a tradução do padrão MPEG-7 em ontologia OWL.	Informação sobre as ontologias multimídia, de fundamentação e de domínio envolvidas no projeto.	Informação sobre o subconjunto de elementos MPEG-7 MDS contemplados na ontologia.
Recursos de conhecimento utilizados	Padrão MPEG-7	Padrão MPEG-7	Padrão MPEG-7; DOLCE; SUMO.	Padrão MPEG-7
Disponibilidade de conhecimento externo	Sítio <i>W3C Multimedia Vocabularies on the Semantic Web</i> .	Sítio <i>W3C Multimedia Vocabularies on the Semantic Web</i> .	Ausente	Ausente
Disponibilidade de testes	Ausente	Ausente	Ausente	Ausente
Disponibilidade de resultados de avaliação de testes	Ausente	Ausente	Ausente	Ausente
Equipe de desenvolvimento	Grupo de pesquisa académico. Ontologia referenciada no <i>W3C Multimedia Vocabularies on the Semantic Web</i> .	<i>Distributed Multimedia Application Group</i> e <i>Music Technology Group</i> (Universitat Pompeu Fabra). Ontologia referenciada no <i>W3C Multimedia Vocabularies on the Semantic Web</i> .	<i>German Research Center for Artificial Intelligence</i>	<i>Pervasive Computing Research Group, department of Informatics and Telecommunications</i> (University of Athens)
Projetos e ontologias que fazem uso	Projetos Harmony, Fuel Cell e Visible Cell.	Projeto ReDeFer e ontologias associadas; Music Brainz ontology.	Projeto SmartWeb	Projeto Polysema

Fonte: elaborado pela autora.

QUADRO 5 - Quadro sinóptico da análise de conteúdo das ontologias multimídia

Aspectos analisados		Media Ontology	COMM	Boemie	M3 Multimedia	M3O	MPEG-7 Hunter	MPEG-7 Rhizomik	SmartWeb MPEG-7	Polysema MPEG-7
Linguagem de implementação		OWL DL	OWL DL	OWL DL	OWL DL	OWL	OWL Full	OWL Full	RDFS	OWL
Clareza no código		Nomenclaturas claras para os conceitos.	Estrutura taxonômica complexa advinda da DOLCE e padrões multimídia.	Nomenclaturas claras para os conceitos.	Nomenclaturas claras para os conceitos.	Entidades organizadas por padrões de projeto de conteúdo de ontologias.	Definições das classes e propriedades sem comentários significativos.	Taxonomia extensa e com ausência de comentários em sua estrutura.	Taxonomia extensa, mas com comentários em sua estrutura.	Taxonomia simples, mas com ausência de comentários em sua estrutura.
Adequação a extração de conhecimento		Subclasses específicas para características de mídia.	Modularização para padrões multimídia.	Modularização para segmentos e descritores de mídia.	Modularização para multimídia, multidomínio e multilingua.	Modularização para padrões multimídia e de proveniência.	Subclasses específicas para características de mídia.	Terminologia análoga ao padrão MPEG-7.	Terminologia com prefixos dos recursos utilizados.	Taxonomia específica para descrição multimídia.
Adequação a convenção de nomes		Padrões de metadados W3C.	Padrão de metadado MPEG-7.	Padrão de metadado MPEG-7.	Padrões de metadados W3C; MPEG-7.	Terminologia adotada pela equipe de projeto.	Padrão de metadado MPEG-7.	Padrão de metadado MPEG-7.	Padrão de metadado MPEG-7.	Padrão de metadado MPEG-7.
Representação de Metadados Multimídia	Independentes de conteúdo	Criação, produção, classificação, informação e uso da mídia.	Informação da mídia.	Local de acesso a mídia e arquivo da mídia.	Criação, produção, classificação, informação e uso da mídia.	Informação da mídia.	Criação e produção da mídia.	Criação, produção, classificação, informação e uso da mídia.	Criação, produção e informação da mídia.	Criação, produção e classificação da mídia.
	Dependentes de conteúdo	Ausente	Características visuais.	Características visuais e de áudio.	Características visuais e de áudio.	Ausente	Características visuais.	Características visuais e de áudio.	Características visuais.	Ausente
	Descritivos de conteúdo	Segmentos de mídia.	Padrões multimídia de decomposição e anotação semântica.	Segmentação multimídia.	Segmentação, semântica de conteúdo e personalização.	Padrões multimídia de decomposição, anotação semântica e coleções.	Segmentos de mídia.	Segmentação, semântica de conteúdo, personalização e áudio de alto nível.	Decomposição de segmentos e semântica de conteúdo multimídia.	Segmentos de mídia.
Características semânticas	Axiomas	Presente *	Presente	Presente	Presente	Presente	Ausente	Presente	Ausente	Presente *
	Anotações	Presente	Presente	Presente	Presente **	Presente	Presente **	Ausente	Presente	Presente **

Fonte: elaborado pela autora.

* - Presença de axiomas com declarações lógicas simples.

** - Presença de anotações, porém sem declarações conceituais consistentes.

4 CONCLUSÕES

A criação e uso de metadados torna-se essencial para a padronização e descrição de recursos de informação, facilitando a recuperação, o uso ou o gerenciamento de fontes de informação digital (ABBAS, 2010). Na inclusão de metadados em documentos multimídia na Web, ontologias e formatos de metadados padrão W3C podem ser endereçados, por exemplo, a organização e ao controle da terminologia usada em metadados; e para auxílio de usuários na produção de anotações mais sistemáticas e consistentes por meio de conhecimento explícito acerca de um domínio (OSSENBRUGGEN, NACK e HARDMAN, 2004).

Conforme assinalam Siqueira e Silva (2011), apesar do conceito abrangente e comumente empregado para metadado, a saber, “dado sobre dado”, os usos, as sintaxes e as apropriações são diferentes em escala, complexidade e custo. Desse modo, a questão central pode ser posta da seguinte forma: como indexar, catalogar e recuperar eficazmente conteúdos multimídia para as inúmeras tipologias de metadados existentes a variadas necessidades e conjunturas. Os resultados de pesquisa evidenciados neste artigo conseguem contribuir na perspectiva de possíveis soluções para o tratamento dos variados tipos de metadados existentes para descrição de acervos com conteúdo multimídia, conforme pode ser conferido nas ponderações que se seguem.

A comparação de várias propostas de ontologias multimídia frente a padrões de metadados ISO como o MPEG-7 e o Dublin Core evidenciou características relevantes (conforme mostra o Quadro 5) que podem e devem ser descritas para melhor recuperação de recursos multimídia, principalmente no contexto da Web. A necessidade de integração semântica e disponibilização global de recursos multimídia na rede é um propósito comum entre as propostas de ontologias pesquisadas.

A partir das constatações das características multimídia concernentes a cada uma das ontologias analisadas, podem-se selecionar os recursos de conhecimento provenientes de suas estruturas a partir do confronto com requisitos determinados em projetos que demandem reuso de conceitos para tal domínio.

Um modelo conceitual pode ser proposto a partir do reuso das ontologias multimídia investigadas nesta pesquisa e vir a ser uma solução abrangente para representação multimídia. As propostas investigadas, geralmente, ou focam em aspectos mais genéricos envolvendo padrões multimídia como anotação e segmentação, mas sem contemplar tipologias específicas de metadados para descrição multimídia; ou oferecem descritores específicos para algumas

categorias de metadados, mas sem raciocínios de modelagem pertinentes que possam assegurar, por exemplo, a extensibilidade da conceituação.

Por fim, uma avaliação metódica, bem fundamentada e criteriosa realizada em ontologias multimídia previamente analisadas pode promover condições necessárias para a seleção e o reuso de recursos de conhecimento apropriados a um determinado domínio. Nesse sentido, tal método seria uma estratégia interessante para selecionar e alinhar eficientemente vocabulários e ontologias multimídia concebidos por comunidades distintas de modo a cobrir satisfatoriamente aspectos ontológicos da realidade documental de tipo multimídia.

REFERÊNCIAS

ABBAS, June. *Structures for organizing knowledge: exploring taxonomies, ontologies, and other schema*. New York: Neal-Schuman Publishers, 2010.

ARNDT, R. et al. *COMM: a core ontology for multimedia annotation*. 2009. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.154.5510>>. Acesso em: 2 out. 2013.

BERNERS-LEE, T; HENDLER, J.; LASSILA, O. The semantic web. *Scientific American*, [S.l.], v. 284, n. 5, p. 34-43, May 2001.

BODOFF, D.; HUNG, P.C.K.; BEN-MENACHEM. M. Web metadata standards: observations and prescriptions. *IEEE Software*, [S.l.], v. 22, n. 1, p. 78-85. Jan./Feb. 2005.

DASIOPOULOU, S. *Multimedia content and descriptor ontologies: final version*. 2008. Disponível em: <https://www.academia.edu/2721370/Multimedia_content_and_descriptor_ontologies-final_version>. Acesso em: 23 de mar. 2014.

DODEBEI, V. Memória e conhecimento: oralidade, visualidade e reprodutibilidade no fluxo da informação. In: ENCONTRO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO, 10., 2009, João Pessoa. *Anais...* João Pessoa: Ancib, 2009.

FERNÁNDEZ, M.; GOMEZ-PÉREZ, A.; JURISTO, H. *Methontology: from ontological art towards ontological engineering*. 1997. Disponível em: <http://oa.upm.es/5484/1/METHONTOLOGY_.pdf>. Acesso em: 20 ago. 2013.

GARCÍA, R.; CELMA, O. semantic integration and retrieval of multimedia metadata. In: INTERNATIONAL WORKSHOP ON KNOWLEDGE MARKUP AND SEMANTIC ANNOTATION, 5th, 2005, Galway. *Proceedings...* Galway, 2005, p. 69–80.

GILLILAND-SWETLANDO, Anne J. *Introduction to metadata: setting the stage*. 2000. Disponível em: <http://www.getty.edu/research/conducting_research/standards/intrometadata/pdf/swetland.pdf>. Acesso em: 12 maio 2005.

GRUBER, T. R. *What is an ontology?* 1993a. Disponível em: <<http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>>. Acesso em: 3 ago. 2013.

GUARINO, N. *Formal ontology in information systems*. 1998. Disponível em: <<http://citeseer.ist.psu.edu/viewdoc/download;jsessionid=E88DA9B5B5A9797C83C1F2E3C907991F?doi=10.1.1.29.1776&rep=rep1&type=pdf>>. Acesso em: 3 set. 2013.

HUNTER, J. Adding multimedia to the semantic web – building an MPEG-7 ontology. In: INTERNATIONAL SEMANTIC WEB WORKING SYMPOSIUM, 1st, 2001, Stanford. *Proceedings...* Disponível em: <https://files.ifi.uzh.ch/ddis/iswc_archive/iswc/ih/SWWS-2001/program/full/paper59a.pdf>. Acesso em: 25 ago. 2013.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS (IFLA). *Functional requirements for bibliographic records*. Munique: K.G. Sauer Verlag, 1998. 144 p. Study Group on the Functional Requirements for Bibliographic Records. Disponível em: <<http://www.ifla.org/files/assets/cataloguing/frbr/frbr.pdf>>. Acesso em: 12 jul. 2013.

INFORMATION STANDARDS QUARTERLY. [S.l.]: NISO, v. 22, n. 1, 2010. Special edition: state of the standards and year in review. Disponível em: <http://www.niso.org/publications/isq/free/FE_DCMI_Harper_isqv22no1.pdf>. Acesso em: 18 jul. 2013.

MARTÍNEZ, J. M. *MPEG-7 overview (version 10)*. 2004. Disponível em: <<http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm>>. Acesso em: 5 dez. 2013.

NACK, F.; OSSENBRUGGEN, J.V.; HARDMAN, L.H. That obscure object of desire: multimedia metadata on the web -part 2. *IEEE MultiMedia*, [S.I.], v.. 12, n. 1 , p. 54-63, 2005.

NIXON, Lyndon et al. Multimedia, broadcasting, and eCulture. In: DOMINGUE, John; FENSEL, Dieter; HENDLER, James A. *Handbook of semantic web technologies*. Heidelberg: Springer-Verlag Berlin, 2011. cap. 21, p.911-975

OSSENBRUGGEN, J. V.; NACK, F.; HARDMAN, L. H. That obscure object of desire: multimedia metadata on the web - part 1. *IEEE MultiMedia*, [S.I.], v. 11, n. 4, p. 38-48, Oct./Dec. 2004.

SAATHOFF, C.; SCHERP, A. Unlocking the semantics of multimedia presentations in the web with the multimedia metadata ontology. In: INTERNATIONAL CONFERENCE ON WORLD WIDE WEB, 19th, 2010, Raleigh. *Proceedings...* New York: ACM, 2010. p. 831-840.

SILVA, Daniela Lucas da ; SOUZA, Renato Rocha . Representação de documentos multimídia: dos metadados às anotações semânticas. In: XIV ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 2013, Florianópolis. *Anais...* Florianópolis: Ancib, 2013.

SIQUEIRA, Ivan C. P.; SILVA, José F. M. da. Metadados: o fio de Ariadne ou a coragem de Teseu? *Bibl. Univ.*, Belo Horizonte, v. 1, n. 1, p. 11-18, jan./jun. 2011. Disponível em: <<https://www.bu.ufmg.br/rbu/index.php/localhost/article/viewFile/35/15>>. Acesso em: 22 set. 2013.

SITARAM, D.; DAN, A. *Multimedia servers: applications, environments and design*. San Francisco: Morgan Kaufmann Publishers, 1999.

SUÁREZ-FIGUEROA, M.C.; ATEMEZING, G.A.; CORCHO, O. The landscape of multimedia ontologies in the last decade. *Multimedia Tools and Applications*, [S.l.], v. 62, n. 2, p. 377-399, Jan. 2013.

SUÁREZ-FIGUEROA, M. C.; GÓMEZ-PÉREZ, A.; FERNÁNDEZ-LÓPEZ, M. The NeOn methodology for ontology engineering. In: SUÁREZ-FIGUEROA, M. C. et al. (Ed.). *Ontology Engineering in a Networked World*. Berlin: Springer, 2012. p. 9-34.

STEGMAIER, F. et al. How to align media metadata schemas? design and implementation of the media ontology. In: INTERNATIONAL CONFERENCE ON SEMANTIC AND DIGITAL MEDIA TECHNOLOGIES, 4th, 2009, Graz. *Proceedings...* [S.l.]: CEUR-WS.org, 2009. Workshop on semantic multimedia database technologies (SeMuDaTe 2009).

VALKANAS, G.; TSETOS, V.; HADJIEFTHYMIADES, S. The polysema MPEG-7 video annotator. In: INTERNATIONAL CONFERENCE ON SEMANTICS AND DIGITAL MEDIA TECHNOLOGIES, 2nd, 2007, Genova. *Proceedings...* Berlin: Springer, 2007. *Demo*.

VEMBU, S. et al. Towards bridging the semantic gap in multimedia annotation and retrieval. In: INTERNATIONAL WORLD WIDE WEB CONFERENCE, 15th, 2006, Edinburgh. *Proceedings...* Edinburgh: [s.n.], 2006. Workshop on Semantic Web Annotations for Multimedia.

ESTUDO COMPARATIVO DE ESTRUTURAS DOCUMENTAIS EM INSTITUIÇÕES ANÁLOGAS: UMA ANÁLISE TIPOLOGICA SOBRE INVENTÁRIOS POST-MORTEM

COMPARATIVE STUDY OF RECORD STRUCTURES AT SIMILAR INSTITUTIONS: DIPLOMATICS ANALYSIS OF POSTMORTEM INVENTORIES

Sonia Maria Troitiño Rodriguez

Resumo: Este estudo se propõe a acompanhar produção documental de duas das mais importantes instituições da justiça brasileira, confrontando-a com as mudanças e alterações na estrutura organizacional das entidades ao longo de dois séculos e meio, período em que coexistiram. Assim, a relação entre função, atividade e documento é recuperada de modo a compreender como o sistema organizacional da instituição interfere na própria criação de registros. Para isso, elegemos a documentação pertencente aos fundos Juízo Ordinário de São Paulo (séc. XVI-XX) e Juízo Ordinário de São Paulo (séc. XVI-XIX), ambos custodiado pelo Arquivo Público do Estado de São Paulo, como parâmetro para essa ponderação. Por serem instituições de longa sobrevivência, com considerável parcela de seus documentos preservados, permitem o estabelecimento de um estudo comparativo da produção de um mesmo tipo documental em diferentes momentos históricos. O tipo documental escolhido é o auto cível de inventário post-mortem, uma das principais fontes utilizadas pelos historiadores e estudiosos para a recuperação de informações sobre o modo de vida da população brasileira durante os períodos de Colônia e Império. O resultado busca mapear as alterações sofridas ao longo do tempo pelos modelos jurídicos adotados para registrar tais ações, assim como apresentar uma proposta de análise tipológica que leve em conta a evolução histórica do tipo documental em ampla escala temporal. Como resultado, ao comparar a produção documental proveniente de distintas organizações, pudemos verificar certa constância na estrutura interna do modelo documental ao longo do tempo. Isso nos fez refletir que, mais do que a organização administrativa ou burocrática que dá origem aos arquivos, o que molda a expressão documental são as regras advindas do direito administrativo, quando em encontro da intencionalidade do registro.

Palavras-chave: Tipologia Documental. Metodologia Arquivística. Inventários post-mortem.

Abstract: This study aims to watch the documents creation from two of the most important institutions of the Brazilian justice system, comparing it with the changes in the organizational structure of the entities over two centuries and a half. Thus, the relationship between function, activity and record is retrieved in order to understand how the organizational system of the institution itself interferes with the creation of records. For this, we chose records from the archives of Juízo de Órfãos de São Paulo (XVI-XX century) and Juízo Ordinário de São Paulo (XVI-XIX century), both belonging to the Arquivo Público do Estado de São Paulo, as a parameter for this analysis. Because they are long-term survival institutions, with a considerable part of their records preserved, allow the establishment of a comparative study of the production of a record type at different historical moments. For this, we chose the postmortem inventories, one of the main sources used by historians and scholars to recover information about the way of life of the population during periods of Brazilian colony and empire. The search results map the changes experienced over time by legal models adopted to register such processes as well as submit a proposal for documentary analysis that takes into account the historical evolution of record type in broad timescale. As a result, when comparing archives from different organizations, we observed certain constancy in the internal structure of the records models over time. It made us reflect that, rather than

administrative or bureaucratic organization that gives rise to archives, which forms the documentary expression are the resulting rules of administrative law, when in meeting the intent of the record.

Keywords: Archival Methodology. Diplomatics. Postmortem inventories.

1 INTRODUÇÃO

O presente trabalho vem expor alguns dos resultados alcançados pelo projeto findo Estudo comparativo de funções e atividades em instituições análogas: a produção de inventários post-mortem nos Juízo Ordinário e Juízo de Órfãos de São Paulo. Este estudo se propôs a acompanhar produção documental de duas das mais importantes instituições da justiça brasileira, confrontando-a com as mudanças e alterações na estrutura organizacional das entidades ao longo da época colonial, período em que coexistiram. Assim, buscamos recuperar a relação entre função, atividade e documento de modo a compreender como o sistema organizacional da instituição interfere na própria criação de registros. Nesse intuito, elegemos a documentação pertencente aos fundos Juízo Ordinário de São Paulo (séc. XVI-XX) e Juízo Ordinário de São Paulo (séc. XVI-XIX), ambos custodiado pelo Arquivo Público do Estado de São Paulo (APESP, Brasil), como parâmetro para essa ponderação. Por serem instituições de longa sobrevivência, com considerável parcela de seus documentos relativamente preservados, permitiram o estabelecimento de um estudo comparativo da produção de um mesmo tipo documental em diferentes momentos históricos.

O tipo documental escolhido é o auto cível de inventário post-mortem, uma das principais fontes utilizadas por historiadores e estudiosos para a recuperação de informações sobre o modo de vida da população, pelo fato de que mesmo sendo um documento considerado “oficial”, o tipo de informação que registra é em grande parte de ordem privada, revelando aspectos do cotidiano como religiosidade, cultura, família, legado, etc. A escolha por esse recorte documental pautou-se, entre outros motivos, no grande volume de documentos preservados - cerca de 4000 processos -, o que conseqüentemente permitiu acesso a dados mais exatos.

Ao longo da pesquisa, pudemos perceber e explorar, o sistema de produção documental dessas entidades, assim como outros temas transversais que insidem contundentemente no sistema de organização arquivística dos acervos em questão. Entre os temas abordados, nos debruçaremos aqui, especialmente, nos relativos à história institucional e tipologia documental, assim como expor uma proposta metodológica de análise documental voltada para documentos do judiciário.

2 CONTEXTUALIZANDO DOCUMENTOS: HISTÓRIA INSTITUCIONAL DOS JUÍZOS ORDINÁRIO E DE ÓRFÃOS DE SÃO PAULO

A recuperação da história institucional permite o revelar de praxes e transformações burocráticas sofridas pelas entidades ao longo de sua existência. Conhecer o sistema organizacional de uma instituição, identificando funções, atividades e procedimentos, leva a compreensão do mecanismo do sistema de registro da informação e, conseqüentemente, à contextualização da produção documental.

O Juízo Ordinário e o Juízo de Órfãos formam dois dos principais pilares da justiça portuguesa, sistema implantado no Brasil enquanto colônia. Em ambos os casos, seus principais cargos eram ocupados por moradores considerados “homens bons”, eleitos anualmente, quando juiz ordinário, ou trienalmente, quando de órfãos, para ocuparem o cargo.

Apesar de configurarem aparelhos de justiça distintos, com representantes próprios, não raramente um mesmo vereador acumulava as funções de juiz ordinário e de órfãos, causando confusão entre cargos e atribuições. Toda essa mescla organizacional aparece refletiva nos processos judiciais decorrentes de suas atividades, levando a necessidade de uma análise documental mais apurada para identificar a proveniência e, assim, o fundo.

Os juízes ordinários ou juízes da terra, como eram conhecidos por serem moradores do município, era a representação mais clara da magistratura de 1ª instância na justiça portuguesa (CARRILLO, s/d). Possuía um campo de atuação bastante amplo, cuja jurisdição abrangia causas de foro civil e criminal no âmbito do município do qual cuidavam, de forma ampla – salvo as exceções previstas na legislação, como nas ações pertencentes ao Juízo de Órfãos. Contudo, quando não houvesse juiz de órfãos na localidade, o juiz ordinário deveria assumir suas atribuições, porém sempre respeitando o regulamento próprio do direito orfanológico. Tinham igualmente a responsabilidade de revisar e controlar a atuação de juízes de instâncias inferiores, vereadores, oficiais de justiça e seus próprios antecessores, dentro de seu perímetro de atuação.

A atuação dos juízes ordinários evocava o direito consuetudinário, funcionando assim como uma justiça local, estruturada sobre o direito costumeiro. Tinha como seu anverso o juiz de fora, magistrado letrado e designado pelo rei para atuar junto às comunidades e representantes dos interesses da Coroa.

Conforme determina as Ordenações Filipinas, no Livro 1, Tit. 65, § 1, os juízes ordinários deveriam sempre levar varas vermelhas, enquanto que os juízes de fora varas brancas. Dessa forma, visualmente, era possível à população reconhecer a face da justiça na

qual estava diante. Como herança desse costume, ainda na atualidade, as divisões internas judiciais presididas por um juiz de direito denominam-se varas.

No Brasil, o conjunto de leis que regravam a atuação da justiça ordinária, pode ser encontradas para o período inicial, século XVI, nas Ordenações Manuelinas (Livro 1, tít. 44.), e a seguir com as Ordenações Filipinas, (Livro 1, tít. 65). Além destas, amparavam-se nas Leis Extravagantes e Leis e Decretos Regionais – considerando o período de existência da instituição.

Em 1832, a Lei de 29 de novembro, promulgou o *Código do Processo Criminal de Primeira Instância com disposição provisória acerca da administração da Justiça Civil*. A nova organização judiciária do império extinguiria o juízo ordinário. Segundo Candido de Almeida, ao detalhar a dissolução dessa instituição e do cargo que a representa, refere-se a sua figura como uma “excrecência” do direito não profissionalizado, cuja presença fazia frente ao poder real. (CANDIDO DE ALMEIDA, 1870, notas).

A atuação correspondente ao juiz ordinário, na nova organização da justiça, ficou a cargo do juiz de paz. Porém suas atribuições foram bastante modificadas. No caso, de ações judiciais, anteriormente cuidadas pelos juízes ordinários, que demandavam aplicação de direito formal, como no dos *inventários post-mortem*, a partir desse momento passaram a fazer parte do rol de atividades dos novos juízes de direito das comarcas.

Paralelamente a ação dos juízes ordinários, entre os séculos XVI e XIX, mas com atuação bastante aproximada, diferenciando-se as causas que lhes correspondiam em torno da condição jurídica dos envolvidos nas ações, estavam os juízes de órfãos.

O processo orfanológico era entendido como aquele no qual se descrevia, avaliava e repartia o patrimônio dos que deixavam por sua morte herdeiros menores ou incapazes da administração de seus bens (CARVALHO, 1879). A competência desses processos pertencia a um juizado especial, o Juízo de Órfãos, definido no título 88 das Ordenações Filipinas. Ao lado dos juízes dos feitos da fazenda, provedor dos resíduos e capelas e defuntos e ausentes, o juiz de órfãos fazia parte dos juízes privativos, que eram instituídos em razão da pessoa, a qual era colocada sob a administração de um juiz de direito (PLÁCIDO E SILVA, 1980). Sendo assim, o juiz de órfãos é o juiz de direito competente ou homem bom, constituído por autoridade legítima, para cuidar das pessoas a incapazes equiparadas, conforme as Ordenações Filipinas, Liv. 1, Tít. 88; e o Alvará de 24/10/1814. Ramalho (1874) chegou a afirmar que nenhum cargo da magistratura tinha maior importância do que o de juiz de órfãos, em razão das atribuições e extensão de sua jurisdição perante a sociedade.

Da implantação do judiciário brasileiro, até a década de 1830, as questões orfanológicas mantiveram um juízo privativo exclusivo para seu atendimento, até que a Lei de 3 de Novembro de 1830 determinou a extinção da Provedoria de Ausente, assim como seu regimento, leis e provisões, passando sua administração e arrecadação pertencer ao Juízo dos Órfãos. Essa mesma lei determinou que o cartório fosse transferido ao escrivão dos órfãos (Suzano, 1854). A partir de então, as atribuições do Ramo de Defuntos e Ausentes, anteriormente órgão do Juízo da Provedoria, passa ser também de competência do Juízo de Órfãos da capital (RAMALHO, 1874; POVEDA, 2007).

Em 1891, a Lei Paulista nº 18, de 21 de novembro, reorganizou a Justiça do Estado de São Paulo e determinou que em cada comarca do estado houvesse um juiz de direito, com exceção da capital, onde foram criadas cinco varas: cível, criminal, órfãos e ausentes, feitos da fazenda e provedoria. Com a reestruturação administrativa da justiça paulista, a antiga composição cartorária do Juízo dos Órfãos e seu respectivo arquivo foram incorporados à nova vara, ficando o Tribunal de Justiça de São Paulo responsável em última instância pela guarda da documentação. (POVEDA, 2007).

Já em 1892, a Lei 80 de 25 de agosto, amplia a estrutura do Juízo de Órfãos e do Ordinário, criando segundas varas em São Paulo, capital, para atender a crescente demanda oriunda do desenvolvimento urbano emergente. O artigo 2º do Decreto nº 3432 de 31 de dezembro de 1921, que regulamenta a execução da lei nº 1759 de 17 de novembro do mesmo ano, que reforma a organização judiciária do Estado de São Paulo, manteve duas Varas de Órfãos e Ausentes na Capital.

Apesar de alguns estados brasileiros manterem até hoje em sua estrutura judiciária varas denominadas “de órfãos”, em São Paulo, o Decreto-Lei 11.058, art. 25, de 26 de abril de 1940, extinguiu as Varas de Órfãos e Ausentes e, em substituição, criou as vara de Família e Sucessões, ficando esta última responsável pelas atribuições e competências do Juízo de Órfãos e herdando sua estrutura administrativa:

Dessa forma, o Juízo de Órfãos da Comarca de São Paulo termina sua história, colocando fim em mais de 350 anos de desenvolvimento de funções e registro de atividades.

3 INVENTÁRIOS POST-MORTEM: CARACTERIZAÇÃO DO TIPO DOCUMENTAL NO EIXO TEMPO-ESPAÇO EM RELAÇÃO À SUA PROVENIÊNCIA

Inventários post-mortem são até hoje um dos tipos documentais mais usuais produzidos pela Justiça. Sua origem, enquanto ação da judicial registrada está ligada ao direito das sucessões e remonta ao Império Romano. Consiste em um processo judicial de

feitura obrigatória responsável pela descrição, avaliação e divisão dos bens que a pessoa possuía na ocasião de sua morte. Assim, essa ação tem por objetivo apurar a real situação econômica do falecido, levando em consideração, além do arrolamento de suas posses, as dívidas ativas e passivas a fim de garantir que cada herdeiro obtenha a parte que lhe é de direito dentro da partilha que se seguirá.

Ainda que nosso estudo de caso verse sobre documentos coloniais e imperiais brasileiros, pudemos notar ao consultar manuais jurídicos de diversos períodos históricos, confrontando-os com processos judiciais contemporâneos, é possível perceber que há uma constância formular que os *inventários post-mortem* ainda mantêm na atualidade¹⁷⁰.

Segundo Macedo Soares (1906), o inventário pode ser entendido como o processo que tem por finalidade o arrolamento, liquidação e partilha os bens possuídos em comum a título de sucessão. Este entendimento do processo de inventário, como sinônimo de inventário post-mortem, é compartilhado por vários autores e constantemente apresentado pelos manuais da época. Apenas encontramos em Gouvêa Pinto (1881), ao tratar da questão da sucessão, uma compreensão mais ampla do sentido do inventário, entendendo-o como qualquer descrição de bens com ou sem avaliação.

Ramalho (1874) divide o inventário em solene e simples ou amigável. Diz-se solene quando é feito pelo juiz com citação de todos os herdeiros e interessados, guardando-se as solenidades de direito, enquanto que o simples se faz unicamente com descrição dos bens perante um tabelião e testemunhas.

As matérias que escapassem a investigação própria do inventário deveriam ser discutidas em diligências paralelas. Contudo, são passíveis de discussão durante o processo de inventário algumas questões que estão diretamente relacionadas ao seu teor, como as relativas à filiação, à qualidade do filho quando legitimado por casamento, à habilitação de herdeiros, à nulidade do testamento, assim como a capacidade da pessoa em exercer a testamentária (PLÁCIDO E SILVA, 1980).

No tratado *Primeiras Linhas Orfanológicas*, Carvalho (1879) detalha o formulário de um inventário com todos os autos, termos, certidões e despachos pertinentes, alertando que as únicas duas diferenças existentes entre o processo de inventário post-mortem processado no Juízo dos Órfãos e os processados em outros juízos, como o Juízo Ordinário, era o fato de estes últimos serem obrigatoriamente requeridos por alguma das partes e não ex-officio como

¹⁷⁰ Hoje em dia regrados no Brasil pelas Leis Federais nº 5869, de 11 de janeiro de 1973 e nº 11.441, de 4 de janeiro de 2007.

os de causa orfanológica, além da característica de não haver nomeação de curador como os tramitados no Juízo de Órfãos.

Uma das principais diferenças entre *inventários post-mortem* produzidos no Juízo Ordinário e no Juízo de Órfãos, diz respeito ao período de tramitação do processo. No Juízo Ordinário, uma vez apurados todos os bens, dívidas ativas e passivas e os legatários envolvidos, o processo tinha o seu fim com a conclusão do arrolamento e divisão de bens na partilha. Já na justiça orfanológica, a finalização da partilha, não era coincidente com a do processo judicial. Uma das principais características desse tipo documental, dentro do Juízo de Órfãos, é de continuidade do processo, que seguirá transcorrendo até a emancipação o menor.

A emancipação configura-se como o momento em que o menor tem acesso à sua legítima herança. Marca a finalização do papel do juiz de órfãos diante da pessoa curatelada, dessa forma tornando toda e qualquer ação subsequente relativa ao não mais incapaz, objeto pertencente à jurisdição de outros juízos: Ordinário, de Direito Civil, Criminal, Municipal, Comercial, de acordo com a matéria demandada e o período histórico em que se encontra.

Contudo, os filhos menores que obtivessem carta de emancipação, mesmo dispensados do pátrio poder, estariam sujeitos por morte dos seus pais à jurisdição dos juízes de órfão, que deveriam fazer o competente inventário e partilha, sem embargo das cartas de emancipação.

Dessa forma e sintetizando, o inventário post-mortem, diferentemente do testamento, era um processo obrigatório que deveria ser realizado sobre o montante de bens da pessoa falecida, independentemente de sua condição social ou da quantia envolvida. Matérias que escapassem da investigação própria do inventário, que era composto basicamente pelo arrolamento, liquidação e partilha de bens, deveriam ser discutidas em diligências paralelas, dando origem a outros processos correlativos.

4 APRESENTAÇÃO DO MÉTODO DE ANÁLISE DOCUMENTAL

Neste trabalho nos propusemos a analisar a fórmula documental dos processos de inventário post-mortem de duas instituições análogas, o Juízo de Órfãos e o Juízo Ordinário, ambos da comarca de São Paulo, e no que diferenciam suas estruturas documentais. Com essa intenção, desenvolvemos uma metodologia específica a ser aplicada para a identificação tipológica para essa documentação, levando em consideração as particularidades dos acervos.

É preciso dizer, que a metodologia que adotamos é em larga medida baseada no tão difundido modelo espanhol para reconhecimento do tipo documental, desenvolvido pelo Grupo de Trabajo de los Archiveros Municipales de Madrid (1988). Desse modo, para o

estabelecimento de uma proposta metodológica de análise tipológica a ser empregada nos documentos de justiça, pautada na tipologia documental, houve influência da originalmente estabelecida pelo grupo de arquivistas espanhóis para a administração municipal.

Uma versão anterior de nossa proposta metodológica para arquivos do judiciário foi apresentada na tese de doutoramento *O Juízo de Órfãos de São Paulo: caracterização de tipos documentais (séc. XVI-XX)* (TROITIÑO, 2010). Contudo, nesse trabalho o formato de análise tipológica apresentada destinava-se exclusivamente para os documentos de uma instituição em específico. O que procuramos realizar aqui foi a ampliação desse modelo de análise, ainda voltado para o Judiciário, porém adaptado para qualquer um de seus aparelhos jurídicos dentro do eixo tempo-espaço.

Em busca da adequação da técnica ao objeto a ser aplicado e tendo em conta que a proposta de análise apresentada por esses arquivistas destina-se a documentos provenientes do âmbito do executivo espanhol e que a primeira versão deste modelo era bastante restritiva, realizamos algumas adaptações criando, assim, uma proposta de análise própria, condizente com as especificidades dos documentos próprios do poder judiciário brasileiro em larga escala.

Desse modo, os originalmente campos *destinatários, duração do trâmite, ordenação, vigência e expurgo*, constantes no modelo espanhol, foram omitidos ou excluídos devido ao fato dessas informações não serem condizentes ao teor informacional da documentação da Justiça. Em contrapartida, houve a necessidade da criação de novos campos descritivos, que expusessem importantes informações acerca desse corpus documental. Nesse intuito, foram estabelecidos os seguintes novos campos a constar no modelo de análise proposto:

TIPO DOCUMENTAL: é definido pela espécie documental, acrescida da atividade que a gerou. Segundo Bellotto (2002, p.96), sua definição pode basear-se “*na legislação, em tratados de direito administrativo, manuais de rotinas burocráticas, glossários, dicionários terminológicos ou a partir do próprio documento*”. No caso deste estudo, especificamente, sempre será inventários post-mortem, porém optamos por mantê-lo justamente para preservar sua abrangência.

TÍTULO ORIGINAL: refere-se ao título dado pelo escrivão no momento da autuação do processo. Apresenta grande variedade na nomenclatura, chegando muitas vezes a ocorrer imprecisão no emprego de termos, quando em relação com o formulário documental. Na primeira versão do modelo para a justiça, o nome do campo era *identificação original*. Contudo, em geral, as normas de descrição arquivística (ISADg, Nobrade, Neda, etc),

possuem o campo denominado *título*, ainda que em outra acepção, muito mais genérica que a aqui proposta, por permitir a inserção de elementos dispares, entre eles a eleição entre o *título original* (nome tradicionalmente empregado nas descrições arquivísticas) e o *atribuído*.

DEFINIÇÃO: procura estabelecer e evidenciar o significado da ação movida, caracterizando-a em relação a função e atividade geradora do documento.

PRINCIPAIS INCIDÊNCIAS: faz menção às principais situações e assuntos tratados pelo tipo documental, procurando apresentar as conjunturas sociais sobre a qual determinada ação é movida. Sofre influência direta da Diplomática, principalmente da apresentada por Duranti (1996) ao relacionar o *ato* e o *fato* e através de sua expressão documental.

TRÂMITES PROCESSUAIS: pode ser entendido como cada um dos estados ou diligências que, dentro de um determinado procedimento ou trâmite legal, deve ser seguido até a finalização do pleito. Procura evidenciar as rotinas e protocolos de trabalho a serem seguidos pelos agentes produtores de documentos em uma organização.

DOCUMENTOS BÁSICOS QUE COMPÕEM OS AUTOS: principais itens documentais, indispensáveis para a composição dos processos, estando assim relacionados com a própria estrutura do documento. Porém, não significa que estes sejam os únicos itens encontrados em determinado processo. Dependendo da natureza da ação e dos objetivos que pretendem ser alcançados, são incorporadas outras peças documentais relacionadas. Na primeira versão do modelo, este campo chamava-se *partes que compõem os autos*, entretanto verificou-se que esta denominação era incoerente por insinuar que todos os tipos documentais de processo produzidos no âmbito do judiciário apresentavam a mesma sequência documental. Em realidade, existem documentos que obrigatoriamente devem constar no processo, mas identificamos uma grande variação em volume e espécies que compõem os autos, resultantes da demanda judicial em questão.

EVOLUÇÃO HISTÓRICA DO TIPO DOCUMENTAL: procura, através do método comparativo entre os processos analisados, rastrear mudanças na composição documental dos tipos ao longo do tempo.

DOCUMENTOS CORRELATIVOS: este campo faz menção a outros tipos documentais que estão analogamente relacionados ao descrito, buscando dessa maneira o cruzamento de informação entre a ação registrada e outras ações correlativas.

Para explicitar melhor os pontos de tangência entre o tradicional modelo desenvolvido pelo grupo de arquivistas de Madri e a proposta apresentada e empregada nesta pesquisa, elaboramos um QUADRO comparativo entre as duas versões:

QUADRO 1 - Comparativo entre modelos de análise tipológica

Modelo de Madrid	Modelo para documentos da Justiça	Comparação entre modelos
Entidade produtora	Entidade produtora	Equivalentes
Tipo documental (<i>denominação, definição, código, características externas</i>)	Tipo documental	O modelo para o judiciário toma como parâmetro a identificação tipológica equivalente à denominação aplicada no modelo de Madrid. Em relação aos subitens, criou-se campo próprio para definição ; os códigos de classificação para os fundos Juízo Ordinário e Juízo dos Órfãos de São Paulo não foram estabelecidos; e as características externas sempre serão: gênero textual, suporte papel, formato maço, forma original
∅	Título original	Discriminação da variedade de títulos encontradas nos processos judiciais
Legislação	Legislação	Equivalentes
∅	Definição	Campo desmembrado de tipo documental
Destinatário	∅	No modelo de Madrid, o destinatário sempre será o <i>Ayuntamiento</i> , na Justiça sempre será a própria entidade da <i>Justiça</i> produtora
Conteúdo	Principais incidências	Equivalentes
Trâmite	Trâmites processuais	Equivalentes
Duração do trâmite	∅	Na Justiça os prazos variam de acordo com cada ação em particular

Documentos básicos que compõem o expediente	Documentos básicos que compõem os autos	Equivalentes
Ordenação da série	Ordenação da série	Equivalentes
Vigência administrativa	∅	Na Justiça a duração dos processos varia de acordo com cada ação, sendo arquivado logo após a finalização da demanda
Expurgo	∅	Não aplicável, por serem considerados documentos históricos (anteriores a 1940) Decreto 60145/2014 e artigo 31 do Decreto nº 48.897/2004. Para processos posteriores, há a Resolução 483/2009 e Recomendação TJSP nº 37/2011 e Provimento TJSP 1743/2010.
∅	Evolução histórica do tipo documental	Rastreia alterações na composição documental dos tipos ao longo do tempo
∅	Documentos correlativos	Cruza informações entre séries relacionadas

Acreditamos que o modelo aqui exposto, possa ser expandido e também usado para a análise de processos cíveis e criminais provenientes de outros órgãos da justiça, em qualquer período histórico. O resultado concreto da aplicação deste método, além da análise crítica sobre a evolução histórica da produção, tramitação e guarda dessa documentação, foi o estabelecimento e análise tipológica individualizada de dois conjuntos documentais, pertencentes a diferentes fundos, mas batizados com o mesmo nome.¹⁷¹

Ainda que em essência a tramitação dos assuntos fosse igual, porque assim o sinalizava a lei, na prática existia diferentes possibilidades para a sua realização e, sobretudo, múltiplas variantes na configuração e formato do tipo documental, com grande diversidade dos documentos básicos (itens documentais) que o integram. Contudo, como a tipologia documental sugere, a busca pelo estabelecimento do tipo documental fundamenta-se no padrão e nunca na exceção. Dessa forma, é o reconhecimento das características comuns (fórmula jurídica, função e atividade geradora) entre as diferentes unidades documentais que possibilita a identificação dos vários tipos documentais.

¹⁷¹ Em anexo pode ser encontrado o modelo de ficha de análise tipológica preenchido com dados extraídos dos Fundos Juízo Ordinário e Juízo de Órfãos de São Paulo.

Heredia (2007) defende que a unidade documental nunca deve ser confundida com o tipo documental. Justamente por essa razão, não há obrigatoriedade do nome do tipo documental, como recurso classificatório arquivístico, ser coincidente ao do dado pelo escrivão no ato do assentamento.

5 CONSIDERAÇÕES FINAIS

Assim como Concepción Mendo Carmona (2004), entendemos ser necessário descender a história das instituições produtoras do documento para poder entender e aplicar efetivamente o princípio da proveniência, sendo que para a correta aplicação do princípio é essencial conhecer a estrutura da instituição produtora, seus métodos de funcionamento, seus procedimentos de trabalho e as variações que sofre ao longo do tempo. Desse modo, a identificação do contexto funcional da produção do documento possibilita vincular a existência material do registro à cadeia de ações que motivaram tal ato, assim como estabelecer a inter-relação deste com as estruturas organizacionais que lhe respaldam. Perante isso, destaca-se a importância da compreensão da tipologia documental como instrumento interpretativo dentro da metodologia empregada para a análise documental.

Estabelecer critérios de análise que levem em conta a origem dos documentos faz-se imprescindível para a contextualização funcional da produção documental. O trabalho de organização de fundos arquivísticos deve necessariamente passar pelo entendimento das causas que motivaram a criação dos documentos, razão de sua própria existência, possibilitando assim a identificação de referenciais a serem utilizados para o agrupamento de informações e documentos nos distintos níveis hierárquicos em que se encontram, dentro de um plano de classificação arquivística capaz de evidenciar a organicidade inerente ao arquivo.

Tendo isso em conta, buscamos neste trabalho conferir em que medida documentos de um mesmo tipo documental, produzidos em momentos históricos distintos, conservaram o padrão de produção documental sobre o qual se apoiavam. Aliamos, ainda, a comparação entre a produção documental proveniente de distintas estruturas administrativas, de um mesmo âmbito, e pudemos verificar certa constância na estrutura interna do modelo documental de um determinado tipo. Isso nos fez refletir que, mais do que a organização administrativa ou burocrática que dão origem aos arquivos, o que molda a expressão documental são as regras advindas do direito administrativo, quando em encontro da intencionalidade do registro.

Pudemos constatar que a caracterização de tipos documentais, através do estudo das diferentes partes que conformam a unidade documental, sua tramitação, relação entre produtor

e acumulador, legislação a qual está subordinada, vigência administrativa, conteúdo, avaliação, prazos de guarda e possíveis mutações do tipo documental ao longo de extensos períodos históricos, trazem a tona informações vitais para o processo de organização arquivística. Todas essas informações traduzidas na composição do documento fazem parte da análise tipológica e corroboram para a gestão documental por reger a produção documental, assim como para pautar o processo classificatório e o controle das distintas fases do ciclo vital, visando à avaliação e destinação de documentos.

O estudo da compatibilidade entre estruturas de composição documental entre diferentes unidades documentais possibilita o seu agrupamento por semelhanças e, conseqüentemente, a seriação. Nesse sentido, a caracterização tipológica da documentação, inquestionavelmente auxilia o estabelecimento de modelos a serem adotados para o reconhecimento de unidades documentais com semelhante origem, uso e tramitação. Quando, então, aliada a contextualização proporcionada pela história institucional, proporciona maior completude no que tange ao valor informacional e ao entendimento sobre produção documental.

Em *Tempo e Circunstância*, Camargo e Goulart, elucidam que

“Em busca da lógica orgânica dos conjuntos documental, não há como escapar do esforço da tipificação que procura amalgamar o conhecimento da estrutura formal dos atos escritos, obedientes a modelos e legitimados do ponto de vista administrativo e jurídico, ao conhecimento das atividades que lhe deram origem e que justificam sua funcionalidade básica: dispor, obrigar, conceder direitos, comunicar, provar, informar e testemunhar” (CAMARGO e GOULART, 2007, p.66).

Nota-se que, independentemente do período temporal em que foram produzidos, em geral documentos de um mesmo tipo preservam igual estrutura interna de composição documental – até porque, mudanças de impacto na estrutura do documento, acabariam por gerar novos tipos documentais. Essa questão ganha força quando pensamos, dentro da arquivística, na eleição do método de classificação funcional como um sistema organizacional capaz de proporcionar contextualização aos documentos, independentemente de balizas temporais impostas.

Essa característica de estabilidade dos modelos pode, inclusive, ser constatada em momentos de transição atravessados pelas instituições produtoras, como os ocorridos na década de 1830, quando o ramo de defuntos e ausentes é incorporado ao Juízo de Órfãos de São Paulo; na década de 1890, quando o judiciário paulista se reorganiza e são criadas duas varas de órfãos; e no ano de 1917, com a implantação do Código Civil Brasileiro em substituição as Ordenações Filipinas. No caso do Juízo Ordinário, ele coexistiu com o Juízo

de Órfãos até 1830, quando foi extinto e em seu lugar o recém criado Juízo de Direito assumiu suas funções relativas ao direito das sucessões.

Em geral, os estudos existentes sobre tipologia documental concentram-se na produção de uma determinada instituição. Falam em modelos documentais, mas sempre se referindo a estudos de casos. Com esta pesquisa, pudemos perceber que o tipo documental possui uma estrutura bastante normalizada, independentemente da origem do documento. No entanto, é importante frisar a importância do princípio de proveniência, pois a partir dele é estabelecido o plano de classificação arquivística. É impossível desvincular conceitualmente a proveniência da seriação de documentos.

Desse modo, ratificamos a posição de correntes, como as representadas por Antonia Heredia Herrera (2013) e Heloísa Bellotto (2002 e 2004), que entendem o tipo documental como modelo, desvinculando-o do registro em si. É o tipo documental componente fulcral na composição do documento, mas nunca o próprio.

REFERÊNCIAS

- ALMEIDA, Cândido Mendes de (Com.). **Código Filipino ou Ordenações e Leis do Reino de Portugal**. 14^a ed. RJ: Typ. do Instituto Philomathico, 1870. 5 vols.
- BELLOTTO, Heloísa Liberalli. **Arquivo Permanente: tratamento documental**. 2. ed. RJ: Editora FGV, 2004.
- BELLOTTO, Heloísa. Como fazer análise diplomática e tipológica de documentos de arquivo. SP: APESP/Imprensa Oficial, 2002.
- CAMARGO, Ana Maria e GOULART, Silvana. Tempo e circunstância: a abordagem contextual dos arquivos pessoais. SP: IFHC, 2007.
- CARRILLO, Carlos Alberto. **Memória da Justiça Brasileira**. Bahia: Tribunal de Justiça do Estado da Bahia, s/d. 3 vol. Disponível em http://www.tj.ba.gov.br/publicacoes/mem_just/mem_just.htm, acessado em 25/01/2014.
- CARVALHO, José Pereira de. **Primeiras Linhas sobre o processo orfanológico**. RJ: AA. da Cruz Coutinho Ed., 1879.
- DURANTI, Luciana. **Diplomática: usos nuevos para uma antigua ciência**. Sevilha, Espanha: S&S Ediciones, 1996.
- FAORO, Raymundo. **Os donos do poder**. RJ: Ed. Globo, 2001. p. 213-214.
- GOUVÊA PINTO, Antonio Joaquim. **Tratado dos Testamentos e Sucessões**. (Ed. comentada por Augusto Teixeira de Freitas). RJ: B.L. Garnier, 1881.
- GRUPO DE TRABALHO de los Archiveros Municipales de Madrid. **Manual de tipologia documental de los municipios**. Madrid: Comunidad de Madrid, 1988.

HEREDIA HERRERA, Antonia. *En torno al tipo*. **Revista Arquivo & Administração**. Rio de Janeiro: AAB, v. 6, n. 2, jul/dez 2007. p.25-50.

HEREDIA HERRERA, Antonia. *En torno al tipo*. **Manual de Archivística Básica: gestión y sistemas**. México: Archivo Histórico BUAP, 2013

MACEDO SOARES, Oscar de. **Manual do Curador Geral dos Órfãos**. 2. ed. RJ: H. Garnier, 1906.

PLÁCIDO E SILVA. **Vocabulário Jurídico**. 6. ed. RJ: Forense, 1980. 5vols.

POVEDA VELASCO, Ignácio M. **Os Esponsais no Direito Luso-Brasileiro**. SP: Quartier Latin, 2007.

RAMALHO, Joaquim Ignácio. **Instituições Orphanologicas**. SP: Typ. de Jorge Seckler, 1874.

SÃO PAULO. Coleção de Leis e Decretos do Estado de São Paulo, vários anos.

TROITIÑO, Sonia. **O Juízo de Órfãos de São Paulo: caracterização de tipos documentais (séc. XVI-XX)**. 2010. Tese (Doutorado em História Social) - Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2010. Disponível em <<http://www.teses.usp.br/teses/disponiveis/8/8138/tde-27042010-103207/>>.

ANEXOS: FICHAS DE ANÁLISE TIPOLÓGICA PREENCHIDAS.

FICHA DE INVENTÁRIO POST-MORTEM DO JUÍZO ORDINÁRIO DE SÃO PAULO

ENTIDADE PRODUTORA	Juízo Ordinário da Comarca de São Paulo
TIPO DOCUMENTAL:	Autos cíveis de inventário post-mortem.
TÍTULO ORIGINAL:	<ul style="list-style-type: none"> Autos de inventário.
LEGISLAÇÃO	Ord. Liv. 1, Tít. 65 e 88; Ord. Liv. 4, Tít. 94-97.
DEFINIÇÃO:	Processo judicial responsável pela descrição, avaliação e divisão dos bens que a pessoa possuía na ocasião de sua morte. Esta ação tinha por objetivo apurar a verdadeira situação econômica do falecido, levando em consideração além do arrolamento de suas posses, as dívidas ativas e passivas para garantir que cada herdeiro obtenha a parte que lhe é de direito dentro da partilha que se seguirá. O inventário post-mortem, diferentemente do testamento, é um processo obrigatório que dever ser realizado sobre o montante de bens da pessoa

falecida, independentemente de sua condição social ou valor envolvido. Matérias que escapem a investigação própria do inventário, que era composto basicamente pelo arrolamento, liquidação e partilha de bens, deveriam ser discutidas em diligências paralelas, que davam origem a outros processos relacionados.

**PRINCIPAIS
INCIDÊNCIAS:**

Processo obrigatório e comum sobre o montante de bens de qualquer pessoa falecida.

TRÂMITES PROCESSUAIS:

O processo iniciava-se a requerimento da parte. Era realizada a autuação e o juramento do inventariante. Caso a pessoa falecida tivesse deixado testamento, este era acostado ao processo com suas respectivas contas. Em seguida lavrava-se o termo de título dos herdeiros, no qual os legatários legítimos à herança eram relacionados. Louvadores eram chamados para realizar a avaliação dos bens, sendo lavrados os termos de louvação e juramento dos louvados. Os bens, divididos em categorias, eram arrolados com seus respectivos valores descritos ao lado, na seguinte ordem: móveis, semoventes, de raiz, dívidas ativas e passivas, gastos com o funeral e bens d'alma, além de qualquer outra informação relacionada a eles. Era feita a alimpação do inventário, que consistia em apresentar a avaliação e os termos da partilha diante dos herdeiros e demais interessados para o esclarecimento de dúvidas e não concordâncias. As partilhas eram realizadas com a descrição dos pagamentos efetuados.

**DOCUMENTOS BÁSICOS
QUE COMPÕEM OS
AUTOS:**

- Autuação
- Juramento do inventariante
- Título dos herdeiros
- Termo de louvação
- Auto de avaliação dos bens (louvação)
- Juramento dos avaliadores
- Relação de bens móveis e imóveis, dívidas ativas e passivas
- Partilhas
- Declarações
- Quitações
- Notificações
- Auto de arrematação
- Sentença
- Termo de Publicação

- Termo de conclusão
- Vistas

ORDENAÇÃO DA SÉRIE Inexistente. O fundo não está organização e os documentos dispersos em diversas coleções – nesses casos a ordenação é cronológica.

EVOLUÇÃO HISTÓRICA DO TIPO DOCUMENTAL: A data limite do período analisado cobre do final do séc. XVI ao primeiro quartel do séc. XIX. Os documentos mais antigos, referentes ao final do século XVI e início do XVII, são processo mais enxutos, com poucos bens arrolados (reflexo da própria condição da população) e que muitas vezes traziam apensados os autos de contas de testamento originais, provindos do Juízo dos Resíduos. A partir do século XVII, cada vez se torna mais raro encontrar esses autos de contas de testamento originais, sendo substituídos por traslados de testamento e a discriminação das contas na parte destinada ao arrolamento de bens. Apesar do corpus documental, abranger quase dois séculos e meio, nota-se pela comparação dos processos uma continuidade da fórmula documental. Evidentemente que cada processo é único, e outros tipos documentais são juntados aos autos de inventário de acordo com o desenrolar da ação (como cartas precatórias, promissórias de dívidas, petições de agravo), no entanto a estrutura básica, que a responsável por definir o tipo do documento como tal, no caso inventário, se manteve estável ao longo do período analisado.

DOCUMENTOS CORRELATIVOS: Como o fundo Juízo Ordinário da Comarca de São Paulo não está organizado, não foi possível identificar com quais outras séries os inventários post-mortem se relaciona.

OBSERVAÇÃO Do ponto de vista de estrutura documental e de tramitação processual, os autos de inventário post-mortem do Juízo Ordinário e Juízo de Órfãos de São Paulo são semelhantes.

FICHA INVENTÁRIO POST-MORTEM DO JUÍZO DE ÓRFÃOS DE SÃO PAULO

ENTIDADE PRODUTORA Juízo de Órfãos de São Paulo

TIPO DOCUMENTAL: Autos cíveis de inventário post-mortem.

TÍTULO ORIGINAL:	<ul style="list-style-type: none">• Autos de inventário.
LEGISLAÇÃO	Ord. Liv. 1, Tít. 88, §4º, 8º e 9º.
DEFINIÇÃO:	<p>Processo judicial responsável pela descrição, avaliação e divisão dos bens que a pessoa possuía na ocasião de sua morte. Assim, esta ação tinha por objetivo apurar a verdadeira situação econômica do falecido, levando em consideração além do arrolamento de suas posses, as dívidas ativas e passivas para garantir que cada herdeiro obtenha a parte que lhe é de direito dentro da partilha que se seguirá. O inventário post-mortem, diferentemente do testamento, é um processo obrigatório que dever ser realizado sobre o montante de bens da pessoa falecida, independentemente de sua condição social ou da quantia envolvida. Matérias que escapem a investigação própria do inventário, que era composto basicamente pelo arrolamento, liquidação e partilha de bens, deveriam ser discutidas em diligências paralelas, que davam origem a outros processos relacionados. No entanto, era possível tratar dentro do inventário, assuntos ligados a questões de filiação, qualidade do filho quando legitimado por casamento, habilitação de herdeiros, nulidade do testamento e capacidade da pessoa em exercer a testamentária.</p>
PRINCIPAIS INCIDÊNCIAS:	Processo obrigatório e comum sobre o montante de bens de qualquer pessoa falecida.
TRÂMITES PROCESSUAIS:	<p>O processo iniciava-se a requerimento da parte ou a ex-offício. Era realizada a autuação e o juramento do inventariante. Caso a pessoa falecida tivesse deixado testamento, este era acostado ao processo com suas respectivas contas. Em seguida lavrava-se o termo de título dos herdeiros, no qual os legatários legítimos à herança eram relacionados. Nomeava-se e registrava-se termo de juramento de um curador para zelar pelos interesses do órfão no inventário. Louvadores eram chamados para realizar a avaliação dos bens, sendo lavrados os termos de louvação e juramento dos louvados. Os bens, divididos em categorias, eram arrolados com seus respectivos valores descritos ao lado, na seguinte ordem: móveis, semoventes, de raiz, dívidas ativas e passivas, gastos com o funeral e bens d'alma, além de qualquer outra informação relacionada a eles. Era feita a alimpação do</p>

inventário, que consistia em apresentar a avaliação e os termos da partilha diante dos herdeiros e demais interessados para o esclarecimento de dúvidas e não concordâncias. O curador geral de órfãos dava vista ao processo e finalmente as partilhas eram realizadas com a descrição dos pagamentos efetuados. Lavrava-se termo de tutoria, no qual estava estipulada a pessoa responsável pela educação e manutenção do órfão, o qual prestava juramento. Tinha início, então, um novo processo, o auto de contas, dentro do próprio inventário por ser entendido como uma continuidade do próprio inventário. No auto de contas deveriam constar todas as despesas com a manutenção do menor, assim como as receitas provenientes de seu pecúlio, durante o período de tutela.

**DOCUMENTOS BÁSICOS
QUE COMPÕEM OS
AUTOS:**

- Autuação
- Juramento do inventariante
- Título dos herdeiros
- Termo de louvação
- Auto de avaliação dos bens (louvação)
- Juramento dos avaliadores
- Relação de bens móveis e imóveis, dívidas ativas e passivas
- Partilhas
- Termo de tutela/curadoria
- Declarações
- Quitações
- Notificações
- Auto de arrematação
- Sentença
- Termo de Publicação
- Termo de conclusão
- Auto de contas da tutela

Podem aparecer outros documentos como termo de escusa ou de remoção de tutor, autos de tutela provisionária, autos de emancipação e/ou autos de habilitação

ORDENAÇÃO DA SÉRIE

Série ordenada cronologicamente

**EVOLUÇÃO HISTÓRICA
DO TIPO DOCUMENTAL:**

A data limite do período analisado é 1578-1810. Os documentos mais antigos, referentes ao final do século XVI e início do XVII, são processo mais enxutos, com poucos bens arrolados (reflexo da própria condição da população) e que muitas vezes traziam apensados os autos de contas de testamento

originais, provindos do Juízo dos Resíduos. A partir do século XVII, cada vez se torna mais raro encontrar esses autos de contas de testamento originais, sendo substituídos por traslados de testamento e a discriminação das contas na parte destinada ao arrolamento de bens. Apesar desse corpus documental, abranger quase três séculos e meio, nota-se pela comparação dos processos uma continuidade da fórmula documental. Evidentemente que cada processo é único, e outros tipos documentais são juntados aos autos de inventário de acordo com o desenrolar da ação (como embargos, tutelas, emancipações), no entanto a estrutura básica, que a responsável por definir o tipo do documento como tal, no caso inventário, se manteve estável ao longo do período analisado.

DOCUMENTOS
CORRELATIVOS:

Relacionam-se praticamente com quase todos os outros processos movidos no Juízo de Órfãos.

OBSERVAÇÃO

Do ponto de vista de estrutura documental e de tramitação processual os processos de inventário post-mortem do Juízo Ordinário e Juízo de Órfãos de São Paulo são semelhantes.

MINERAÇÃO DE TEXTOS PARA A EXTRAÇÃO AUTOMÁTICA DE SINTAGMAS NOMINAIS: TENDÊNCIAS E ESTRATÉGIAS PARA RECUPERAÇÃO DA INFORMAÇÃO JURÍDICA

MINING OF TEXTS FOR AUTOMATIC EXTRACTION PHRASES NOMINAL: TRENDS AND STRATEGIES FOR RECOVERY OF LEGAL INFORMATION

Fanny do Couto Ribeiro de Lima

Resumo: Esta pesquisa se utiliza das técnicas de mineração de textos para a extração automática de sintagmas nominais, com vistas a possibilitar uma melhor recuperação da informação jurídica. A pesquisa apresenta estudos recentes desse novo modelo de extração de informação. Com o objetivo de propor as técnicas de mineração de textos, para a extração automática de sintagmas nominais, como tendências e estratégias para recuperação da informação jurídica, neste estudo se toma como amostra um conjunto de 200.000 normas legislativas do estado de São Paulo como forma de aplicar as técnicas propostas e indicar resultados satisfatórios, capazes de capacitar os profissionais no tratamento e representação do conhecimento. Apoiados nos resultados levantados nesta pesquisa, conclui-se que os estudos sobre mineração e sintagmas nominais no Brasil ainda é uma questão incipiente. Para avançar esses limites, considera-se necessário o desenvolvimento de novas propostas integradoras do nível teórico-metodológico nos campos da mineração de textos, de forma que estabeleçam um meio de representação eficaz para a representação de informações em linguagem natural.

Palavras-chave: Mineração de Textos. Sintagmas Nominais. Informação Jurídica. Extração Automática.

Abstract: This research uses the techniques of text mining for the automatic extraction of noun phrases, in order to allow better recovery of legal information. The research presents recent studies of this new model of information extraction. In order to propose the techniques of text mining for the automatic extraction of noun phrases, such as trends and strategies for recovery of legal information, this study sample is taken as a set of 200,000 legislative norms of the state of São Paulo as a way of apply the proposed techniques and indicate satisfactory, able to train professionals in the treatment and representation of knowledge results. Building on the results collected in this study, it is concluded that the study of mining and noun phrases in Brazil is still an emerging issue. To advance these limits, it is considered necessary the development of new proposals integrative theoretical and methodological level in the field of text mining in order to establish an effective means of representation for the representation of information in natural language.

Keywords: Text Mining. Nominal phrases. Legal Information. Automatic extraction.

1 INTRODUÇÃO

A humanidade experimenta uma época de grandes transformações na organização, tratamento e acesso à informação. Neste novo ciclo, o avanço das tecnologias digitais e o uso intensivo de tecnologias de informação surgem como instrumentos que, além de propiciar um novo sistema de comunicação, possibilitam a criação de novos e diversos suportes informacionais. Este momento caracteriza uma sociedade cada vez mais baseada na cooperação e convergência de tecnologias, onde a expansão das informações registradas em

meio digital prossegue sua dinâmica em grande velocidade, permitido o acúmulo de dados em uma velocidade maior que a capacidade humana de processá-los.

Não obstante os progressos permitidos pelo avanço tecnológico no armazenamento e fornecimento de grandes volumes de informação, este cenário faz com que pessoas e organizações tenham entre suas principais dificuldades o gerenciamento eficiente e específico destas informações, principalmente as não estruturadas.

O grande volume de informação registrada em meio digital e o desenvolvimento acelerado de novos estoques de informação cobram o desenvolvimento de ferramentas específicas, e é neste ambiente renovado pelas tecnologias, com o avanço da internet e uma disponibilidade cada vez maior de informação, que a extração automática de informações surge como ferramenta capaz de assegurar a identificação das informações potencialmente relevantes de um texto, através de um conjunto de processos e técnicas que garante acesso e interpretação destas informações, visto que, na sua grande maioria, estas informações não apresentam regularidade quanto à formatação, estrutura e conteúdo.

Já em 2005, Souza, em sua tese de doutorado, investigava o potencial de uso dos sintagmas nominais em processos de indexação automática, e afirmava que estes possuem maior grau de informação semântica, o que o torna mais eficaz como descritor utilizado em processos automatizados de representação de documentos. Este contexto que marca a sociedade moderna faz com que métodos e técnicas de tratamento automático de textos se tornem cada vez mais importante para as atividades relacionadas ao tratamento da informação.

A extração de informação, por objetivar a estruturação e representação de informações em linguagem natural, utiliza-se das técnicas de mineração de textos para representar, sobre tendências ou padrões, grandes volumes de documentos textuais, de forma que represente o conhecimento embutido nessas informações. Sendo assim, este trabalho foi desenvolvido por meio do recorte de uma pesquisa maior no âmbito de um mestrado profissional que estuda a extração de informação semi-estruturada em diários eletrônicos da justiça.

2 MINERAÇÃO DE TEXTO: DESCOBRINDO CONHECIMENTO DE BASES DE DADOS TEXTUAIS

Apesar da sociedade moderna se mostrar cada vez mais dependente da informação como insumo para conduzir suas ações, essa sobrecarga – que, muitas vezes, não segue critérios de organização quanto à sua estrutura e localização – tem modificando o perfil das pessoas e organizações, principalmente no que diz respeito a gerenciamento e recuperação

dessas informações. Sendo assim, o contexto atual é marcado, cada vez mais, pelo desenvolvimento de ferramentas e estratégias específicas e especializadas que buscam assegurar o gerenciamento e acesso às informações realmente relevantes ao interesse de pessoas e organizações que as procuram.

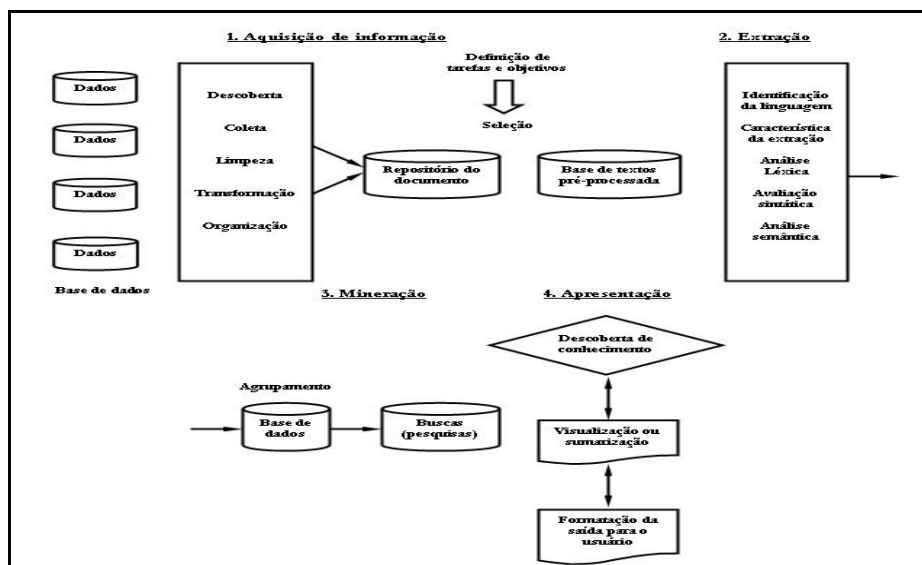
A abordagem da *mineração de texto*, ao aplicar na prática os aspectos da Inteligência Artificial, pode ser vista como uma atividade que comporta uma série de métodos e técnicas capazes de suportar o processo de aquisição do conhecimento em fontes não estruturadas ou textos escritos em formato livre, permitindo, assim, a estruturação e tabulação dessas informações e a descoberta de conhecimento em base de dados textuais. Para Trybula (1999 apud TARAPANOFF; ARAÚJO JÚNIOR, 2006, p. 237), bases textuais são:

[...] coleções de documentos em linguagem natural, sem formato predefinido para seus conteúdos, como acontece com as bases de dados. Dividem-se em bases textuais cujo conteúdo é estruturado de acordo com a sua localização no documento. Como exemplos, têm-se relatórios policiais, relatórios de instituições financeiras, ou seja, o conteúdo pode variar, mas a estrutura do documento é predefinida, e informações não estruturadas, em que se têm como exemplos os relatórios, publicações e a maioria dos documentos textuais.

Para que dados textuais brutos possam se tornar úteis é necessário que eles sejam representados de maneira adequada para a extração de padrões, de regularidades e de tendências de grandes volumes de textos em linguagem natural. Para este objetivo, os processos de *Mineração de Textos* se apresentam como ferramenta capaz de possibilitar a extração não trivial de conhecimentos úteis de dados não estruturados ou semi-estruturados, integrados a grandes bases textuais, previamente desconhecidas, e potencialmente úteis. (PASSOS; ARANHA, 2006).

Com base nos estudos de Ebecken, Lopes e Costa (2005) e Trybula (1999), entende-se que o processo de mineração de textos é dividido em quatro etapas principais, que são: *identificação e aquisição do corpus*, *pré-processamento*, *extração de conhecimento* e *pós-processamento*, detalhado nos diagramas desenvolvidos pelos autores supracitados, nas FIG. 1 e 2.

FIGURA 1: Processo de mineração de textos



Fonte: TRYBULA, 1999.

FIGURA 2: Etapas da mineração de textos

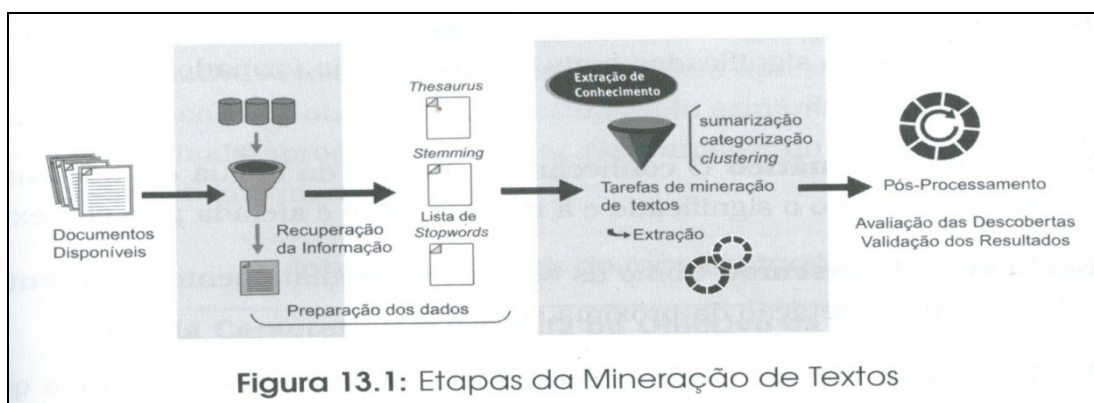


Figura 13.1: Etapas da Mineração de Textos

Fonte: EBECKEN; LOPES; COSTA, 2005.

Em um processo de mineração, primeiramente é necessário identificar, selecionar e coletar uma coleção de textos com a qual se vai trabalhar, chamada aqui de *corpus*. Uma vez selecionado o *corpus* a ser trabalhado, a base de textos passa por um pré-processamento que estrutura e formata a coleção de textos, por meio da transformação da coleção em um formato útil aos algoritmos de aprendizado de máquina, de forma que este seja efetivamente empregado.

Frequentemente, o pré-processamento é caracterizado por excesso de tarefas manuais, envolvendo operações como tratamento e limpeza dos textos, preparando-os para as fases posteriores de execução das tarefas de processamento, assim como das análises dos resultados. Sendo assim, compreende-se que, neste passo, resultam a boa ou má qualidade dos

dados analisados e, conseqüentemente, a eficiência dos algoritmos de aprendizado de máquina utilizados e grande parte da confiança nos resultados obtidos (WEISS et al., 2005).

Na etapa do processamento dos dados textuais, aqui representada pela extração do conhecimento, existem várias tarefas que podem fazer parte do processo de extração, sendo cada uma específica para um tipo de informação. Entre estas tarefas da mineração de textos, podemos citar: agrupamento ou *clustering*; categorização; sumarização, indexação temática, entre outras. (EBECKEN; LOPES; COSTA, 2005).

O pós-processamento compreende a última etapa da mineração de textos, que constitui a validação da descoberta de conhecimento realizada pelo processamento/extração, como também são avaliados os resultados obtidos nas etapas anteriores. A consolidação dessa etapa se apoia no uso de “métricas de avaliação de resultados, ferramentas de visualização e conhecimentos de especialistas” (EBECKEN; LOPES; COSTA, 2005, p. 362) como alternativas indispensáveis para o progresso da pesquisa.

A mineração de textos configura-se – enquanto um conjunto de etapas relacionadas à busca pela representação adequada de conhecimento útil – como um desafio estratégico que se desenrola permanentemente no seu objetivo de extrair conhecimento de ciências como a Linguística e a Informática. A partir daí, é possível entender a Mineração de Textos como “um conjunto de técnicas e processos que descobrem conhecimento inovador nos textos.” (EBECKEN; LOPES; COSTA, 2005, p. 338).

3 EXTRAÇÃO AUTOMÁTICA DE SINTAGMAS NOMINAIS

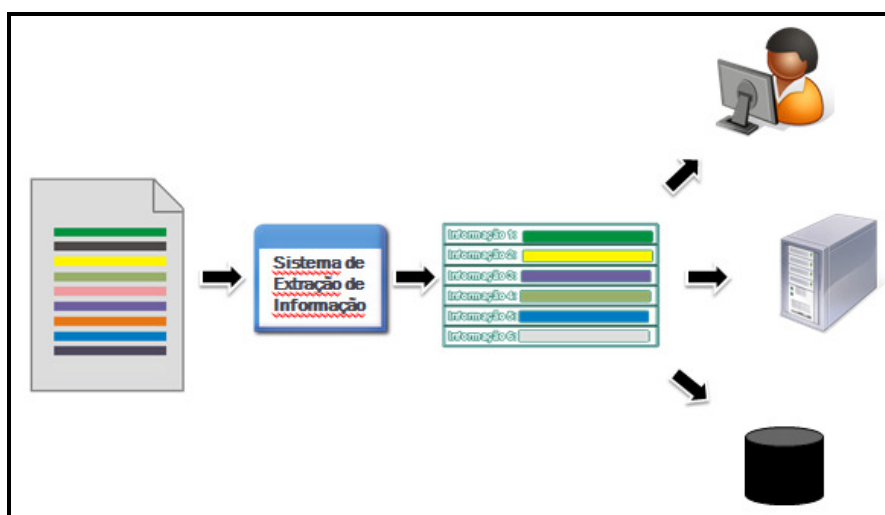
Com o aumento do volume de informações que vem marcando a sociedade contemporânea, métodos e técnicas de tratamento automático de textos se tornam cada vez mais importantes para as atividades relacionadas ao tratamento da informação. Este contexto marca cada vez mais o desenvolvimento de ferramentas e estratégias específicas que buscam assegurar o gerenciamento e acesso às informações realmente relevantes ao interesse de pessoas e organizações. Neste ambiente, a Extração de Informação propõe a extração de informações relevantes de um texto, segundo algum contexto, o que a torna uma das principais alternativas ao aprimoramento e eficiência nos resultados oferecidos pela Recuperação da Informação, já que “através da integração e sintetização da informação, evita a ocorrência de redundâncias em textos que tratam do mesmo assunto” (BATRES et al, 2005, p. 74).

Silva (2003, p. 20), na mesma perspectiva, afirma que:

A extração de informação identifica dados relevantes a um tema contido em um texto e os extrai convertendo-os para uma estrutura tabular. Esta estrutura tem o objetivo de sumarizar o conteúdo do assunto abordado no documento numa forma legível, tanto para o usuário quanto para uma aplicação.

Autores como Wives e Loh (1999 apud BATRES et al, 2005) entendem que a extração de informação, ao invés de encontrar textos que contenham informações e permitir ao usuário procurar o que lhe interessa, preocupa-se em encontrar as informações dentro dos textos e as tratar de forma a apresentar algum tipo de conhecimento novo e útil para o usuário. Porém, apesar do grande volume de informação disponível hoje, estas não apresentam regularidade quanto à formatação, estrutura e conteúdo. Sendo assim, a extração de informação tem, como um de seus principais desafios, a estruturação e representação das informações em linguagem natural, conforme representado nas FIG. 3 e 4.

FIGURA 3: Processo de extração da informação



Fonte: BORGES *et al.*, 2011.

FIGURA 4: A terrorist report and a template of extracted information.

19 March - A bomb went off this morning near a power tower in San Salvador leaving a large part of the city without energy, but no casualties have been reported. According to uno_cial sources, the bomb | allegedly detonated by urban guerrilla commandos | blew up a power tower in the northwestern part of San Salvador at 0650 (1250 GMT).

INCIDENT TYPE	bombing
DATE	March 19
LOCATION	El Salvador: San Salvador (city)
PERPETRATOR	urban guerrilla commandos
PHYSICAL TARGET	power tower
HUMAN TARGET	-
EFFECT ON PHYSICAL TARGET	destroyed
EFFECT ON HUMAN TARGET	no injury or death
INSTRUMENT	bomb

Fonte: GRISHMAN, 1997.

Corroborados pelas afirmações de Jacobs e Rau (1993), de que a extração de informação envolve a identificação de padrões que representam um contexto específico dentro do texto, podemos entender a extração de informação como um conjunto de padrões e filtros que representam e estruturam as informações, contribuindo, dessa forma, para uma melhor recuperação dessas informações.

Como forma de oferecer suporte a esses sistemas e os auxiliar na representação e compartilhamento de conhecimento, por meio da captura e registro da semântica dos elementos terminológicos de um determinado texto, a extração automática de sintagmas nominais surge como uma das principais ferramentas para representação e recuperação adequada de conhecimento útil e relevante.

Para Silva e Morellato (2003; 2007), sintagmas consistem num conjunto de elementos que constituem uma unidade significativa dentro da oração e que mantêm entre si relações de dependência e de ordem, ou seja, sintagmas constituem um conjunto de palavras que, juntas, desempenham uma função na frase, podendo ter função de sujeito ou objeto, dependendo da sua relação com os outros constituintes. Morellato (2007, p. 20) apresenta a regra de formação geral de um sintagma nominal, onde o núcleo é sempre um nome, ou pronome substantivo ou elemento substantivado, que pode constituir um sintagma com apenas uma palavra, ou pode vir acompanhado de outras palavras com as quais estabelecem relações de concordância, que neste caso são representadas por: determinantes (*Det*) e modificadores (*Mod*).

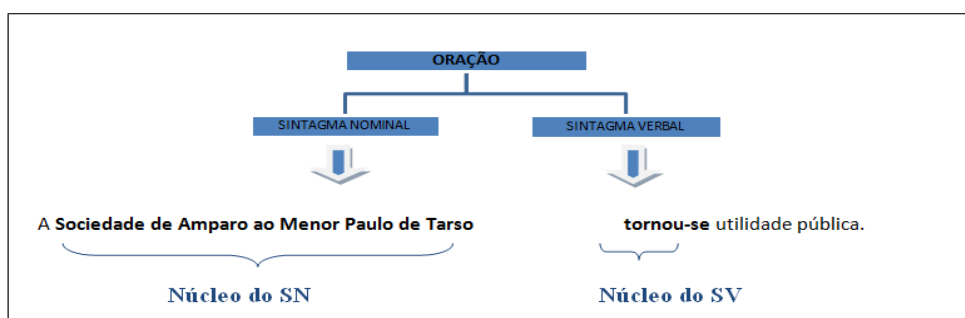
SN – (Det) + (Mod) + Núcleo + (Mod)

A partir da regra apresentada por Morellato (2007), é possível observar que as frases expressam seus conteúdos por meio dos elementos e das combinações desses elementos que a língua proporciona, formando unidades sintáticas dentro de uma unidade maior que é a frase. Quanto aos sintagmas, de maneira reduzida, estes podem ser divididos em sintagmas nominais e verbais. Para que os sintagmas se apresentem como descritores mais eficientes que

as palavras-chave normalmente utilizadas, os sintagmas nominais apresentam sua relevância ao possibilitar maior valor semântico em sua estrutura.

A extração automática de sintagmas nominais traz suas contribuições ao reconhecer e recuperar os sintagmas nominais contidos em frases. Souza (2005), ao estudar o potencial de uso dos sintagmas nominais em processos de indexação automática, afirma que estes possibilitam maior grau de informação semântica e, assim, podem ser mais eficazes como descritores em processos automatizados de representação de documentos, já que analisam a estrutura sintática das frases e retornam apenas seus sintagmas nominais, desconsiderando os demais elementos.

FIGURA 5: Identificação de sintagmas em uma oração



Fonte: AUTOR, 2014.

Uma vez identificados conceitos e características de um sintagma nominal, é possível entender como acontece o processo de extração automática destes sintagmas em um texto em linguagem natural, de forma que possam ser aplicados na criação de termos em sistemas de indexação e busca em documentos, gerando resultados melhores na recuperação de informação.

A extração automática de sintagmas nominais compreende um conjunto de processos e etapas que, segundo Teline, Almeida e Aluísio (2003), possuem duas abordagens: manual e automática. A abordagem manual consiste na obtenção de uma lista de referência e conceitos pré-definidos pelo especialista do domínio. A automática é feita através da ferramenta capaz de gerar listas de N-grama, ou seja, de percorrer um documento e extrair “N” palavras de um determinado texto.

Para este trabalho, adota-se a abordagem automática para a extração de sintagmas nominais, e esta se utiliza das etapas da mineração de textos para realizar a *extração de conhecimento*. Sendo assim, uma vez coletado o *corpus* e realizado o seu pré-processamento, o texto é anotado gramaticalmente, ou seja, são realizadas anotações linguísticas nos textos,

através de etiquetas (tags) que indicam características de uma determinada palavra ou estrutura frasal.

Nos estudos de Baségio (2006), a extração de conceitos também é realizada com base em regras expressas por sequências de etiquetas que, quando encontradas no texto, podem representar termos compostos. A Figura 6 apresenta as regras utilizadas por Baségio:

Figura 6: Regras para extração de termos compostos

1	_SU	_AJ	_PR	_AD	_SU	_AJ
2	_SU	_AJ	_PR	_AD	_SU	
3	_SU	_PR	_AD	_SU	_AJ	
4	_SU	_PR	_AD	_SU		
5	_SU	_AJ	_PR	_SU	_AJ	
6	_SU	_AJ	_PR	_SU		
7	_SU	_PR	_SU	_AJ		
8	_SU	_PR	_SU			
9	SU	AJ				

Fonte: BASÉGIO, 2006.

A partir dos textos anotados linguisticamente, é possível a análise da estrutura sintática das frases e o reconhecimento e extração dos sintagmas nominais, desconsiderando os demais elementos que são considerados irrelevantes (verbos, advérbios, *stopwords*). Visto que a extração dos termos é baseada no método de classe gramatical, faz-se necessário a construção de uma Gramática/Dicionário composta/o de palavras encontradas nos arquivos textos. Corroborando este método, estudos de Baségio (2006) assumem, como ponto de partida para a extração automática de sintagmas nominais, um texto anotado linguisticamente, com as informações associadas a cada palavra do documento: a palavra no seu formato original; o lema da palavra original; a etiqueta gramatical da palavra (exemplo: substantivo, adjetivo, etc.).

O processo utilizado na obtenção da gramática/dicionário de termos aborda diversas etapas, tais como a criação de listas gramaticais (verbos, advérbios, *stopwords*, etc.), que possibilitam a anotação linguística dos textos. Uma vez anotados e extraídos, os sintagmas passam por um processo de refinamento, onde é construído um conjunto de heurísticas que refinam os sintagmas nominais, possibilitando a seleção das classes gramaticais específicas que deseja extrair dos textos, assim como a exclusão de sintagmas que apresentam palavras desnecessárias, consideradas lixo.

Como pode ser visto, a identificação e extração automática de sintagmas nominais é uma tarefa que exige um grande esforço linguístico para a geração das regras de formação

destes sintagmas, para só então ser possível sua recuperação por meio de sistemas inteligentes.

A maioria dos modelos de indexação anteriores considera a utilização de palavras como meio de acesso à informação. Porém, trabalhos de autores como Lopes et al. (2009), Kuramoto (2002), Morellato (2007), Baségio (2006) e Souza (2005) mostram a inadequação do uso das palavras-chave, ou seja, das palavras isoladas para o processo de recuperação de informação, propondo, em seu lugar, o uso de sintagmas nominais, que são considerados mais eficientes como conceitos, pois quando extraídos de um texto, seus significados permanecem os mesmos.

4 INFORMAÇÃO JURÍDICA: UM ESTUDO SOBRE RECUPERAÇÃO E ACESSO

A sociedade contemporânea mostra-se cada vez mais dependente da informação como insumo para conduzir suas ações, tornando-se cada vez mais info-dependente. Os avanços da Tecnologia da Informação reforçam as mudanças econômicas e sociais e nos relatam um ritmo exponencial do crescimento e da influência que a informação tem provocado no perfil das pessoas e organizações.

Ao tratar a questão da informação, Barreto (1994) afirma que a mesma tem se desenvolvido à margem das revoluções e do crescimento industrial, e que essa crescente produção de informação precisa ser armazenada e transmitida de forma eficiente. Neste mesmo enfoque, Capurro e Hjørland (2007, p. 173) sustentam ideia semelhante ao afirmarem que “a produção, distribuição e acesso à informação estão no centro da nova economia.”. Estas circunstâncias estão revolucionando os negócios e a sociedade. A nova espécie de economia – baseada na informação – está emergindo onde o comércio e o investimento são globais e empresas competem com conhecimento, redes digitais e agilidade.

Questão semelhante é levantada por Kuramoto (2006), que traz uma visão inovadora e econômica que enxerga a informação como insumo básico para o desenvolvimento de um país. O autor também enfatiza uma das principais características da sociedade contemporânea – também chamada Sociedade da Informação: a informação ganha *status* de capital como recurso intangível e insumo básico para a construção do conhecimento.

No que diz respeito à informação jurídica, o grande volume de informações se deve ao fato da pesquisa ser um dos principais instrumentos dos advogados, juristas, magistrados e de todos que lidam com a matéria jurídica. Corroborando a ideia de que a informação precisa ser armazenada e transmitida de forma eficiente, Passos (1994, p. 364) defende que “(...) toda essa massa documental de informação jurídica seria apenas o ‘caos documental jurídico’, (*sic*)

se não existissem mecanismos para controlar e recuperar a informação contida nos documentos”.

Ainda nesta perspectiva, é possível observar que a organização da informação jurídica se torna um ponto primordial para a resolução do “caos documental jurídico”, possibilitando a recuperação rápida e eficiente das informações e com isso atrair benefícios como qualidade, produtividade e competitividade para todos que lidam com a matéria jurídica. Neste sentido, Rezende (1998, p. 72) afirma que “não basta, portanto, ter disponíveis as informações, em alguns casos isto pode até dificultar ainda mais os resultados. É preciso que as informações disponibilizadas estejam organizadas e acessíveis a tempo de vencer os prazos e a concorrência”.

Existem diversos conceitos para informação jurídica, ATIENZA (1979 apud REZENDE, 2003, p. 3) afirma que a documentação jurídica é “a reunião, análise e indexação da doutrina, da legislação, da jurisprudência e de todos os documentos oficiais relativos a atos normativos ou administrativos”.

Para este trabalho, o corpus escolhido foram as fontes formais legislativas, as quais são representadas pelas normas superiores (leis, projetos de lei, decretos, medidas provisórias, etc.). As fontes de informação jurídica possuem características que atendem a formatos e estruturas textuais particulares, demandando um conhecimento especializado da linguagem da área do Direito, assim como da estrutura textual dos documentos jurídicos.

Diante das características peculiares da informação jurídica, Passos (1994) aponta também a questão da longevidade deste tipo de informação como objeto de pesquisa, pois mesmo já “ultrapassados” continuam como fontes importantes, a título de consulta.

A informação jurídica tem uma validade quase permanente, interessando ao jurista não apenas a última que foi publicada, a que foi legislada, mas também a anterior, ainda que a revogação de uma norma legal deixe sem efeito essas disposições, seus preceitos serão aplicáveis àqueles atos jurídicos que tenham nascido sob a sua vigência, e portanto (*sic*) a eficácia perdura além da própria revogação. Daí que não se pode prescindir de documentos jurídicos sob o pretexto da antiguidade. (LÓPEZ-MUÑIZ *apud* PASSOS, 1994, p. 364)

Ainda neste contexto, Martinho (2006) caracteriza a informação jurídica a partir dos seguintes itens:

Grandes volumes de informação e rapidez da sua desatualização, face a um constante crescimento e criação de novas fontes (novas leis, novas áreas do Direito, constante produção de doutrinas);

Grande diversidade de fontes de informação;

Proliferação de suportes de informação (bases de dados, CD-ROMs, WWW);

Interação com outras áreas do conhecimento (filosofia, sociologia, economia, história, informática, política, *sic* etc.);

Público-alvo exigente e diversificado (pluralidade de finalidades da informação pretendida);

Necessidade de grande rigor e precisão da sua conservação no documento jurídico, qualquer que seja o seu suporte;

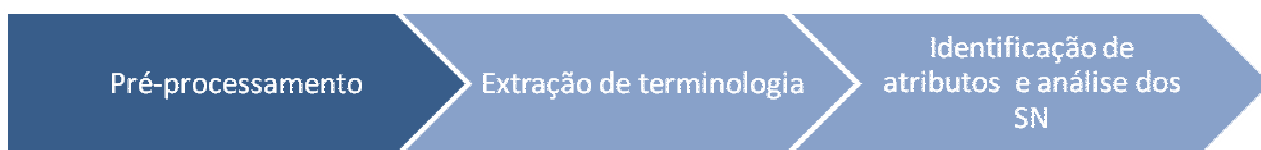
Necessidade de grande rapidez e rigor na sua transmissão, de modo a assegurar a sua correta utilização e aplicação, de acordo com as últimas alterações ocorridas e a garantir a fiabilidade da mensagem transmitida. (MARTINHO, 2006, p. 91).

5 METODOLOGIA PROPOSTA

O trabalho empreendido é um estudo sobre o uso de técnicas de mineração de textos para a extração de sintagmas nominais, métodos que se configuram como tendências e estratégias para recuperação e acesso à informação jurídica. Para tanto, na tentativa de apresentar a relevância desses métodos para a recuperação precisa, apresenta-se aqui a metodologia desenvolvida para auxiliar no mapeamento de informações contidas em normas legislativas (informações não estruturadas), assim como para a extração de padrões, usando, entre outros, dicionário semântico e algoritmos de aprendizado, empregando um conjunto de filtros capaz de representar e recuperar a informação de forma estruturada, a fim de facilitar automaticamente sua manipulação e análise.

A partir daí, a metodologia é composta por três etapas principais, conforme apresentado na FIGURA 7.

FIGURA 7: Etapas da Metodologia



Fonte: AUTOR, 2014.

Para o processo de mineração de textos utilizado nesta pesquisa, a etapa de pré-processamento de textos tem por objetivo montar e formatar a base a ser extraída e analisada; dessa forma, suas principais atividades são:

QUADRO 1: Etapas de Pré-processamento

Etapa de “Pré-processamento”: Construção do Conjunto de Frases Únicas
Etapa de “Pré-processamento”: Tentativa de utilização do segmentador SENTER para a Construção do Conjunto de Frases Únicas
Etapa de “Pré-processamento”: Desenvolvimento de um algoritmo para a “Construção do

Conjunto de Frases Únicas”
Etapa de “Pré-processamento”: Desenvolvimento de um Toquenizador para o algoritmo
Etapa de “Pré-processamento”: Desenvolvimento de um Segmentador para o algoritmo
Etapa de “Pré-processamento”: Realização de testes no algoritmo para a “Construção do Conjunto de Frases Únicas”
Etapa de “Pré-processamento”: Análise estatística com o algoritmo
Etapa de “Pré-processamento”: Realização de testes no algoritmo para avaliar sua capacidade de fornecer estatísticas
Etapa de “Pré-processamento”: Estudo da Base de dados Legiscidade
Etapa de “Pré-processamento”: Elaboração de uma lista de <i>stopwords</i>
Etapa de “Pré-processamento”: Remoção de <i>stopwords</i> no algoritmo
Etapa de “Pré-processamento”: Realização de testes no algoritmo para a construção de frases únicas sem <i>stopwords</i>
Etapa de “Pré-processamento”: Comparação dos resultados com <i>stopwords</i> e sem <i>stopwords</i>
Etapa de “Pré-processamento”: Transformação para minúsculo
Etapa de “Pré-processamento”: Correção ortográfica (em estudo)
Etapa de “Pré-processamento”: Normalização/Aplicação de substituições
Etapa de “Pré-processamento”: Realização de estudos no tesauro do STJ como vocabulário regular
Etapa de “Pré-processamento”: Normalização/Aplicação do lematizador
Etapa de “Pré-processamento”: Realização de testes no etiquetador VISL
Etapa de “Pré-processamento”: Realização de testes no etiquetador Qtag
Etapa de “Pré-processamento”: Realização de testes no etiquetador OGMA
Etapa de “Pré-processamento”: Tentativa de utilização do etiquetador morfossintático TreeTagger
Etapa de “Pré-processamento”: Criação de um etiquetador para o algoritmo
Etapa de “Pré-processamento”: Alteração e composição de um dicionário para o etiquetador

Fonte: AUTOR, 2014.

A extração de terminologia apoia as atividades realizadas pelos responsáveis pelo desenvolvimento e indexação da massa documental que irá compor a base de dados jurídica proposta como objetivo futuro desta pesquisa. Nesta etapa foram realizadas as seguintes atividades:

QUADRO 2: Etapas de Extração de Terminologia

<p>Etapa de Extração de Terminologia: etiquetação dos documentos</p> <p style="text-align: center;">Construção de um dicionário/gramática</p> <p style="text-align: center;">Análise manual dos documentos etiquetados</p> <p style="text-align: center;">Correção do dicionário</p> <p style="text-align: center;">Construção de um algoritmo para a etiquetação dos documentos</p>
<p>Etapa de Extração de Terminologia: correção das etiquetas geradas pelo algoritmo</p>
<p>Etapa de Extração de Terminologia: definição de máscaras (<i>SUB SUB ADJ</i>)</p> <p style="text-align: center;">Pesquisa dos termos frequentes no domínio jurídico</p> <p style="text-align: center;">Construção de uma lista de máscaras</p> <p style="text-align: center;">Correção da lista de máscaras</p> <p style="text-align: center;">Construção de um algoritmo para extração das máscaras</p> <p style="text-align: center;">Análise das máscaras extraídas</p> <p style="text-align: center;">Correção na lista de máscaras</p> <p style="text-align: center;">Correção do algoritmo para extração das máscaras</p> <p style="text-align: center;">Construção de um algoritmo para extração dos descritores</p> <p style="text-align: center;">Análise dos descritores extraídos</p> <p style="text-align: center;">Correção do algoritmo para extração dos descritores</p> <p style="text-align: center;">Construção de um algoritmo para criação das listas de termos e as frequências</p> <p style="text-align: center;">Análise das listas de termos e as respectivas frequências</p> <p style="text-align: center;">Correção do algoritmo para criar as listas de termos e as respectivas frequências</p>
<p>Etapa de Extração de Terminologia: realização de testes na ferramenta de extração Extphrj</p>
<p>Etapa de Extração de Terminologia: realização de testes na ferramenta de extração OGMA</p>
<p>Etapa de Extração de Terminologia: aplicação de heurísticas</p> <p style="text-align: center;">Construção de um conjunto de possíveis heurísticas</p> <p style="text-align: center;">Validação e correção das heurísticas</p> <p style="text-align: center;">Reunião com profissionais da área</p> <p style="text-align: center;">Filtragem sobre as listas de termos que foram geradas</p>

Construção de um algoritmo para aplicação das heurísticas

Análise das listas de termos já corrigidos

Criação de uma lista de termos candidatos

Fonte: AUTOR, 2014.

A primeira atividade na Etapa de Extração de Terminologia compreendeu a etiquetagem dos documentos, para qual foi necessária a construção de um dicionário que associasse a cada termo uma sintaxe pré-definida. Devido à especialidade e às características próprias da linguagem jurídica, as etiquetas utilizadas compreenderam um conjunto de palavras definidas pelo profissional da informação responsável pela qualidade da representação do conhecimento adotada.

Durante a extração de SN, uma vez etiquetados, o sistema utiliza um conjunto de heurísticas para refinar o processo, de forma que sejam extraídos apenas os descritores relevantes ao texto, identificados por um especialista da área de informação. As heurísticas aplicadas aos termos identificados se apresentam da seguinte forma:

QUADRO 3: Heurísticas

<p>1º QUANDO COMEÇAR COM...</p> <p>Nº</p> <p>nº</p> <p>PAULO</p> <p>MESES DO ANO (JAN, FEV, MAR...)</p> <p>TÓPICOS { a); b); c); d)...</p>
<p>2º QUANDO TERMINAR COM...</p> <p>Nº</p> <p>nº</p> <p>ASSEMBLEIA</p> <p>PROJETO</p> <p>SÃO</p> <p>ESTADO</p>

...
<p>3° QUANDO COMEÇAR...</p> <p style="text-align: center;">(SUB + ,)</p> <p style="text-align: center;">(SUB SUB + ,)</p>
<p>4° QUANDO TERMINAR...</p> <p style="text-align: center;">(+ SUB)</p> <p style="text-align: center;">(+ ADJ)</p> <p style="text-align: center;">(+ SUB SUB)</p> <p style="text-align: center;">(+ SUB ADJ)</p>
<p>5° QUANDO UM <i>NGRAMA</i> MENOR TIVER DENTRO DE UM <i>NGRAMA</i> MAIOR</p>

Fonte: AUTOR, 2014.

Em seguida, os sintagmas candidatos a conceitos extraídos são salvos em ordem de frequência, o que permitirá uma melhor análise para a aplicação de ponto de corte.

Quanto à etapa de identificação de atributos e análise dos SN, esta viabiliza uma busca rápida por um documento específico por meio da extração de informação por algoritmos inteligentes e interpretação do conhecimento contido nos textos. Antes da realização de suas principais atividades, para esta pesquisa foi adotada a inserção de um módulo de pós-processamento antes da etapa de análise, conforme é demonstrado no Quadro 4.

QUADRO 4: Etapas de pós-processamento e análise e validação dos SN

Este modelo de mineração de textos propõe o pré-processamento antes da fase de indexação do processo tradicional descrito na etapa de pós-processamento: ordenar os SN nas planilhas por meio da frequência de ocorrência.
Etapa de pós-processamento: descartar os SN que apresentam frequências de ocorrência inferiores.
Etapa de pós-processamento: agrupar os SN remanescentes a partir dos determinantes em suas formas “canônicas” e os reordenar.
Etapa de pós-processamento: analisar manualmente os SN pré-escolhidos e decidir sobre a sua relevância com os descritores, para fins de construção de um a <i>stoplist</i> .
Etapa de pós-processamento: verificar a incidência dos SN nos outros documentos do corpus.
Etapa de pós-processamento: analisar a estrutura e o nível dos SN.
Etapa de pós-processamento: atribuir pontuação e ranquear os SN remanescentes de acordo

com fórmula estabelecida.

Etapa de pós-processamento: caso ocorra empates nos valores da pontuação dos SN, considerar a ocorrência no tesouro do STJ com o fator de desempate.

Etapa de pós-processamento: caso ainda ocorra empates nos valores da pontuação dos SN, considerar os seguintes critérios de desempate:

- a. Maior valor absoluto da frequência de ocorrência;
- b. Menor valor absoluto da ocorrência no corpus;
- c. Maiores nível e estrutura do SN;
- d. Maior quantidade de letras do SN.

Etapa de pós-processamento: apresentar tantos descritores quanto forem desejáveis, a partir da lista ranqueada de candidatos a descritores.

Etapa de análise e validação dos SN: analisar comparativamente.

Etapa de análise e validação dos SN: análise por especialistas.

Fonte: AUTOR, 2014.

No QUADRO 5 são apresentados alguns exemplos da extração de terminologia usando a abordagem linguística. Na primeira coluna estão colocados os padrões sintáticos e as máscaras desenvolvidas para a extração dos sintagmas nominais; na segunda coluna os SN que foram identificados; e na terceira coluna é apresentada a frequência de cada sintagma no texto analisado.

QUADRO 5: Extração de terminologia usando a abordagem híbrida (linguística e estatística)

MÁSCARAS	EXTRAÇÃO DO SN	FREQUÊNCIA DOS TERMOS
SUB SUB SUB SUB SUB	Presidência República Casa Civil Subchefia	Frequência: 1
SUB SUB SUB SUB ADJ	Casa Civil Subchefia Assuntos Jurídicos	Frequência: 1
SUB SUB NUM NUM SUB NUM	Decreto Lei 10.436, 24 abril 2002	Frequência: 1

Fonte: AUTOR, 2014.

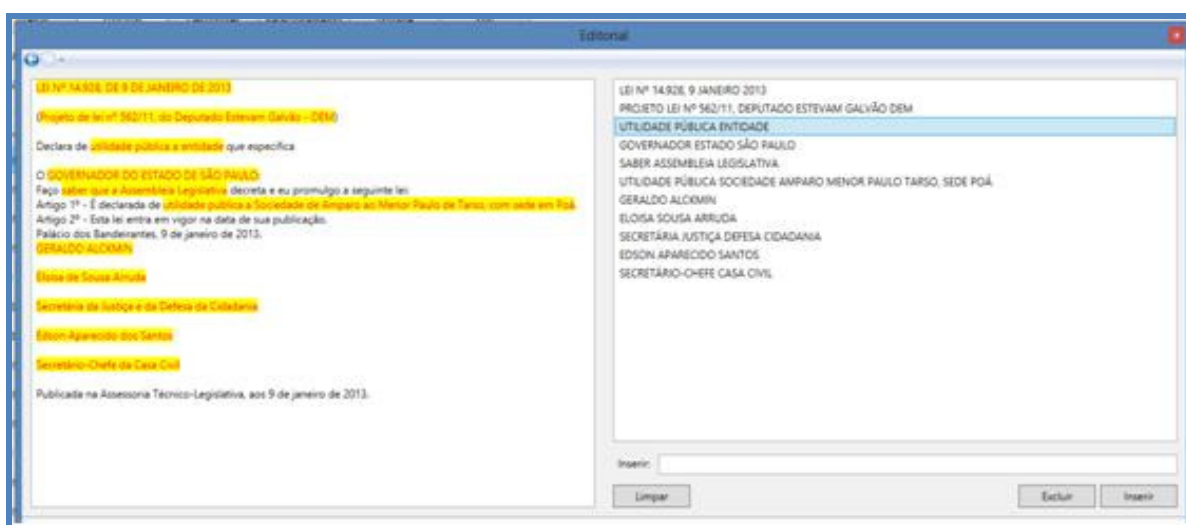
Esta pesquisa propõe a utilização e desenvolvimento de uma metodologia computacional para identificação de sintagmas nominais da língua portuguesa em documentos digitais escritos em linguagem natural, para representação e recuperação da informação jurídica. Sendo assim, depois explicitada a metodologia e os processos adotados para

identificar e extrair sintagmas nominais, a pesquisa teve como resultado o desenvolvimento de uma ferramenta, um sistema identificador de sintagmas nominais da língua portuguesa, voltado para o âmbito da informação jurídica.

A partir do estudo sobre o uso de técnicas de mineração de textos para a extração automática de sintagmas nominais, foi possível desenvolver e implantar uma ferramenta de extração de sintagmas nominais que possibilitou realizar mais facilmente todo processo de indexação automática pensado, inicialmente, para a construção de uma base de dados jurídica destinada à extração e recuperação de termos representativos do conteúdo de documentos textuais que compõem as 200.000 normas legislativas escolhidas como corpus para esta pesquisa.

Conforme a FIG. 8, neste trabalho é apresentado um sistema baseado em extração automática de sintagmas nominais para apoiar as atividades realizadas pelos responsáveis pelo desenvolvimento e indexação da massa documental que irá compor a base de dados jurídica proposta como objetivo futuro desta pesquisa.

FIGURA 8: Sistema de extração automática de sintagmas nominais



FONTE: AUTOR, 2014.

O sistema desenvolvido é composto por um conjunto de módulos e funcionalidades, capaz de interpretar textos em linguagem natural, por meio de análises morfológicas e sintáticas, e recuperar os sintagmas nominais dos textos trabalhados. Segundo Araújo (1995), este tipo de sistema, voltado para a recuperação de informação, é aquele que, entre outras funções, objetiva dar acesso às informações contidas em documentos neles registrados. Os documentos, nesse sistema, contêm informação potencial e são formalmente organizados, processados e recuperados com a finalidade de maximizar o uso da informação.

REFERÊNCIAS

- ATIENZA, C. A. **Documentação jurídica**: introdução a análise e indexação de atos legais. Rio de Janeiro: Achiamé, 1979.
- BARRETO, A. A. Questão da informação. **São Paulo em Perspectiva**, v. 8, n. 4, 1994.
- BASÉGIO, T. L. Uma abordagem semi-automática para identificação de estruturas ontológicas a partir de textos na língua portuguesa do Brasil, Dissertação (Mestrado). Pontifícia Universidade Católica do Rio Grande do Sul – PUCRS, 2006.
- BATRES, E. J. Q. et al. Uso de ontologias para a extração de informações em atos jurídicos em uma instituição pública. In.: **Enc. Bibli: R. Eletr. Bibliotecon. Ci. Inf.**, Florianópolis, n. 19, 1º sem. 2005.
- BORGES, C. **Extração da Informação**. Recife: UFPE, 2011. Disponível em: <www.cin.ufpe.br/~if796/aulas/extracao-2011.pptx>. Acesso em: 15 fev. 2013.
- CAPURRO, R; HJORLAND, B. O conceito de informação. **Perspectiva em Ciência da Informação**, Belo Horizonte, v. 12, n. 1, p. 148-207, jan./abr. 2007.
- EBECKEN, N. F. F.; LOPES, M. C. S.; COSTA, M. C. A. Mineração de textos. In: REZENDE, S. O. (Coord.) **Sistemas inteligentes**: fundamentos e aplicações. São Paulo: Manole, 2005, p. 337-370.
- GRISHMAN, R. **Information extraction**: techniques and challenges. New York: Computer Science Department University New York, USA.
- KURAMOTO, H. Sintagmas nominais: uma nova proposta para a recuperação da informação. **DataGramZero**, v. 3, n. 1, fev. 2002.
- KURAMOTO, H. Informação científica: proposta de um novo modelo para o Brasil. **Ciência da Informação**, v. 35, n. 2, mai./ago. 2006, p. 91-102.
- LOPES, L., et al. ExATO lp - an automatic tool for term extraction from portuguese language corpora. In: LTC'09 - 4th Language and Technology Conference, 2009, Poznan, 2009, Poznan. **Proceedings of the Fourth Language and Technology Conference**. Poznan : Adam Mickiewicz University, 2009, p. 427-431.
- MARTINHO, A. M. O bibliotecário jurídico: identidade e competências profissionais. In: Encontro Nacional de Bibliotecas Jurídicas, 1., 2004. **Anais...** Lisboa: Faculdade de Direito da Universidade de Lisboa, 2006.
- MORELLATO, L. V. **SIDSN**: Sistema Identificador de Sintagmas Nominais. Vitória, 2007.
- PASSOS, E. J. L. O controle da informação jurídica no Brasil: a contribuição do Senado Federal. **Ciência da Informação**, Brasília, v. 23, n. 3, p. 363-368, set./dez. 1994.
- REZENDE, M. I. S. D. Gerenciamento da informação e do conhecimento no escritório de advocacia. In: CIBERÉTICA. SIMPÓSIO INTERNACIONAL DE PROPRIEDADE INTELLECTUAL, INFORMAÇÃO E ÉTICA, 1, 1998, Florianópolis. **Anais...** Florianópolis, SC, 19 a 21 nov. 1998.
- REZENDE, M. I. S. D. Gerenciamento da informação e do conhecimento jurídico. [2003]

SILVA, T. S. M. Extração de informação para busca semântica na web baseada em ontologias. Florianópolis: UFSC, 2003.

SOUZA, R. R. **Uma proposta de metodologia para escolha automática de descritores utilizando sintagmas nominais**. 2005. Tese (Doutorado em Ciência da Informação) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2005.

TARAPANOFF, K.; ARAÚJO JÚNIOR, R. H. Precisão no processo de busca e recuperação da informação: uso da mineração de textos. *Ci. Inf [online]*. 2006, v. 35, n. 3, p. 236-247.

TELINE, M. F.; ALMEIDA, G. M. B.; ALUÍSIO, S. M. Extração manual e automática de terminologia: comparando abordagens e critérios. In: **I Workshop em Tecnologia da Informação e da Linguagem Humana (I TIL)**. ICMC-USP, São Carlos, São Paulo, Out. 2003.

TRYBULA, W. J. Text mining. **Annual Review of Information Science and Technology**, v. 34, p. 385-419, 1999.

WEISS, S. M., et al. **Text mining** – predictive methods for analyzing unstructured information. New York: Springer, 2005.

WIVES, L. K.; LOH, S. Tecnologias de descoberta de conhecimento em informações textuais; ênfase em agrupamento de informações. In: OFICINA DE INTELIGÊNCIA ARTIFICIAL (OIA) III, 1999, Pelotas (RS). **Proceedings...** Pelotas: EDUCAT, 1999, p. 28-48.

CLASSIFICAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO EM CT&I NO CONTEXTO DA DIVULGAÇÃO CIENTÍFICA EM FONTES DE INFORMAÇÃO DIGITAIS

CLASSIFICATION AND REPRESENTATION OF KNOWLEDGE IN CT&I IN THE CONTEXT OF SCIENTIFIC DISSEMINATION IN DIGITAL SOURCES

Arlete Sandra Mariano Alves Baubier
Rosali Fernandez de Souza
Lena Vania Ribeiro Pinheiro

Resumo: Propõe uma reflexão de cunho teórico acerca da organização do conhecimento no âmbito da divulgação científica em fontes de informação digitais, especialmente os periódicos eletrônicos. O enfoque é direcionado para a necessidade de classificação e representação de domínios do conhecimento em Ciência, Tecnologia e Inovação (CT&I), em gêneros textuais do jornalismo científico, a partir da análise de revistas de divulgação da informação em CT&I de instituições de fomento à pesquisa, em especial da revista *Amazonas Faz Ciência*, da Fundação de Amparo à Pesquisa do Estado do Amazonas (Fapeam). O objetivo central é abordar os aspectos históricos e conceituais que permeiam o tema proposto, além de examinar as principais características estruturais que fazem parte desse tipo de publicação. O percurso metodológico do estudo caracteriza-se pela abordagem teórica e descritiva, com aporte bibliográfico e documental. Os principais resultados indicam: a revista encontra-se no processo contínuo de aperfeiçoamento da sua estrutura, procurando adequar-se aos padrões de editoração de periódico de divulgação da informação em CT&I, oferecendo produtos inovadores para públicos específicos e agregando valor e qualidade à publicação.

Palavras-chave: Organização e Representação do Conhecimento. Classificação. Divulgação Científica. Fontes de informação digitais. Ciência, Tecnologia e Inovação (CT&I).

Abstract: This paper proposes a reflection on the nature of theoretical knowledge organization in scientific dissemination of digital information sources, especially electronic journals. The focus is directed to the need for classification and representation of knowledge domains in Science, Technology and Innovation (ST&I), textual genres of science journalism, from the analysis of journals in disseminating information on STI for development institutions to research in particular the magazine *Amazonas Makes Science*, Foundation for Research Support of the State of Amazonas (Fapeam). The main objective is to address the historical and conceptual aspects that permeate the theme, in addition to examining the main structural features that are part of this type of publication. The methodological approach of the study is characterized by theoretical and descriptive approach, within put bibliographic and documentary. The main results indicate: the magazine is in the continuous process of improving its structure, tries to adapt to the standards of publishing periodical dissemination of information on ST&I, offering innovative products for specific audiences and adding value and quality to the publication.

Keywords: Organization and Representation of Knowledge. Classification. Scientific Dissemination. Sources of digital information. Science, Technology and Innovation.

1 INTRODUÇÃO

Nas últimas décadas, o desenvolvimento da comunicação e divulgação científicas foi impulsionado pelo processo de globalização e pelo avanço vertiginoso das diferentes

tecnologias de informação e comunicação (TICs), associadas às diversas fontes de informação e canais de comunicação existentes que influenciam diretamente na nova ordem mundial.

No cenário globalizado, o uso de recursos tecnológicos de acesso à informação e ao conhecimento é cada vez maior, e intenso também é o volume de dados produzidos e disponibilizados em meio eletrônico que necessitam de tratamento do conteúdo semântico. Em atendimento às necessidades informacionais da comunidade científica, que se interessa pelos mais variados assuntos em diferentes áreas do conhecimento da Ciência, Tecnologia e Inovação (CT&I), prioritárias para o desenvolvimento do país, a organização do conhecimento se faz necessária.

Entende-se que estando o conhecimento organizado, haverá melhores condições para a obtenção das informações relevantes quanto às tendências, perspectivas e desafios contemporâneos em áreas estratégicas da CT&I, que se impõem ao campo científico.

A Organização do Conhecimento é área tradicional de pesquisa e ensino em Ciência da Informação. Grande parte da literatura nesta área trata do processo, das atividades e dos instrumentos especialmente desenvolvidos no tratamento de documentos para armazenamento, disseminação, recuperação e uso em sistemas e serviços de informação (SOUZA, 2006, p. 27).

De modo semelhante, Langridge (2006, p. 19) postula que a Organização do Conhecimento, no sentido mais amplo da expressão, pode ser compreendida como aquela que abrange o processo de classificação do conhecimento e como tal constitui uma das principais funções desempenhadas pela biblioteca. Ainda segundo o autor, “É possível haver unidades isoladas de *informação*, mas o termo *conhecimento* implica organização” (LANGRIDGE, 2006, p. 19, grifo do autor).

Desde a Antiguidade, o ato de ordenar, organizar, classificar e representar diferentes seres, elementos, produtos e objetos, entre outros materiais da natureza, era um processo praticado por diferentes civilizações, ainda que não tivesse a denominação pela qual conhecemos hoje. Esse processo mental é inerente à natureza humana. Consciente e inconscientemente, o homem tende a classificar tudo que o cerca, por meio de características e atributos comuns dos objetos, observando as similitudes e as diferenças existentes.

A função que a classificação desempenha na organização do conhecimento é das mais importantes, mesmo imprescindível; concorre efetivamente para que sistemas de recuperação da informação alcancem seu principal objetivo: satisfazer as necessidades de informação e documentação de seus usuários (FERNANDEZ, 2006, p. 3).

No princípio, o pensamento científico tendia para uma lógica de organização do conhecimento de base filosófica, que iniciava a partir da concepção do Universo e tudo aquilo que faz parte dele. Mais tarde, a partir do século XVII, também surgiram às classificações

bibliográficas, quando houve a necessidade de organizar os conteúdos dos documentos de cada campo do saber científico.

Em outras áreas do conhecimento científico, este processo de classificação também se faz presente. Para a Ciência da Informação (CI) interessa destacar a “Organização e Representação do Conhecimento”, buscando os aportes históricos e conceituais na Documentação e Biblioteconomia, onde a gênese desse subcampo da CI encontra sedimentação teórica em seus princípios basilares.

Este artigo propõe uma reflexão de cunho teórico acerca da organização do conhecimento no âmbito da divulgação científica em fontes de informação digitais, especialmente os periódicos eletrônicos. O enfoque é direcionado para a necessidade de classificação e representação de domínios do conhecimento em Ciência, Tecnologia e Inovação (CT&I), em gêneros textuais do jornalismo científico, a partir da análise de revistas de divulgação da informação em CT&I de instituições de fomento à pesquisa, em especial da Fundação de Amparo à Pesquisa do Estado do Amazonas (Fapeam). O objetivo central é abordar os aspectos históricos e conceituais que permeiam o tema proposto, além de examinar as principais características estruturais e pré-textuais que fazem parte desse tipo de publicação.

2 DA COMUNICAÇÃO À DIVULGAÇÃO CIENTÍFICA: APONTAMENTOS HISTÓRICOS E CONCEITUAIS

Ao refletir sobre o tratamento da informação, tomam-se por base a literatura científica em CI, especialmente a bibliografia nas subáreas de Organização e Representação do Conhecimento, Comunicação e Divulgação Científicas. Verificam-se nas contribuições teóricas, as origens, os principais termos e conceitos fundamentais na Ciência da Informação, dentre os quais destacamos os seguintes: noção de ciência, documento, periódico eletrônico e recuperação da informação.

Braga explica que:

O reconhecimento da Ciência como poderosa força motivadora da civilização, afetando suas estruturas intelectuais, econômicas e políticas, trouxe a necessidade de analisar e compreender sua natureza (BRAGA, 1974, p. 156).

Fato é que a Ciência possui uma natureza interdisciplinar e, por isso “A Ciência da Ciência é bem mais vasta que o termo que a designa, abrangendo várias Ciências da Ciência e talvez várias Humanidades da Ciência”. Por sua natureza essencialmente diversa e independente de toda e qualquer atividade humana, a Ciência possui vida e ordem próprias, “intransigente às vontades humanas, impermeável a nacionalismos e enquadramentos

filosóficos, alheia aos desejos e temores da sociedade [...]” (BRAGA, 1974, p. 155; 158), sendo “regularmente reinventada e reescrita através da constante busca e aquisição de novos conhecimentos” (RIBEIRO; PINHEIRO; OLIVEIRA, 2007, p. 2). Nesse sentido, a História e a Filosofia da Ciência, como domínios de conhecimento da Sociologia da Ciência, nos leva a compreender que não há apenas um método a ser seguido no campo científico, mas sim vários métodos que ajudam a compreender o seu próprio estatuto científico.

Historicamente, mais precisamente no período medieval, o hábito de trocar cartas e reminiscências era a forma mais comum de comunicação, ou melhor, “a única forma tradicional de publicação”. Personalidades científicas como Arquimedes e Dositheos, Ptolomeu e Ciro, Peter Peregrinus e Sygerius de Foncaucourt utilizavam estes tipos de correspondências como forma de registrar conversas, decisões, acordos e tratados (BRAGA, 1974, p. 160).

Corroborando com exposto, Hernandez Cañadas (1987, p.19) explica que a origem da literatura científica, bem como das academias e sociedades científicas ocorreram a partir da troca de correspondências entre os cientistas. De igual modo, Oliveira, Valério, Chalhub e Pinheiro (2010, p. 4) ressaltam que a publicação periódica, na forma como conhecemos hoje, resulta de “um longo processo de evolução que começou com a troca de correspondência entre colegas, buscando opinião para suas investigações e idéias, antes de lançá-las para um grupo maior”. Assim, no decorrer dos tempos, outros canais de comunicação da informação científica foram surgindo e se consolidando, como os periódicos científicos.

No século XVII, houve o surgimento das primeiras publicações periódicas de cunho científico de que se tem notícia, *Le Journal des Sçavans*, denominado posteriormente *Journal des Sçavants* e o *Philosophical Transactions of the Royal Society of London* (PINHEIRO, 2005, p. 23), iniciado por Henry Oldenburg, em 1665. Segundo Stumpf (1996, p. 2), estes periódicos são considerados marcos históricos da comunicação da ciência porque “[...] contribuíram como modelos distintos para publicação científica”.

No Brasil, a história do periodismo científico revela que o surgimento das primeiras publicações desta natureza remonta ao século XIX, com a fundação da imprensa no país e quando houve a criação de muitas instituições científicas, pela Corte Portuguesa, destinadas à prática e ao estudo das ciências. Segundo Pinheiro *et. al.* (2009, p. 3), “[...] as primeiras notícias sobre ciência divulgadas em jornais datam de 1808, ano da chegada da família real portuguesa e da fundação da Imprensa Régia [...]”.

Miranda e Pereira (1996, p. 375) revelam que o periódico científico é um “veículo formal de comunicação” que tem vital importância no processo de comunicação científica,

pois visa atender basicamente duas vertentes: “a de comunicação do conhecimento e a de comunicação entre os pares da comunidade científica”. Meadows (1999, p. 7), ao estudar com profundidade a comunicação científica, também define que o periódico científico configura-se como um tipo de canal formal da ciência, sendo de suma importância para o processo de comunicação porque viabiliza a legitimação da pesquisa científica. Além disso, observa-se que uma publicação têm, no mínimo, duas funções essenciais: “**memória** – arquivo do conhecimento – e a função **disseminação** – transmissão de amplas idéias” (BRAGA; OBERHOFER, 1982 apud KRZYZANOWSKI; FERREIRA, 1998, p. 166, grifo da autora).

Pode-se conjecturar que esta relação **documento – memória – informação** foi defendida anteriormente por Paul Otlet, um dos mais célebres documentalistas da história da Documentação, considerado o “pai da Documentação” e um dos principais precursores da gênese da Ciência da Informação. Otlet era um visionário que buscava integrar ao conceito de documento outros objetos e mídias, isto é, outras formas de transmissão da informação e desejava também “expandir o conceito de biblioteca, de tal forma, que ela fosse um lugar onde pessoas acessassem qualquer tipo de informação sabendo não apenas qual é o documento, mas o que ele contém” (PEREIRA, 2000, p. 2).

Naquele tempo, Otlet já propunha outras vertentes para o entendimento do amplo significado do termo documento, inclusive associado a algo que constitui a memória materializada da humanidade. Ou seja, para Otlet (1934, p. 43), o documento é um “receptáculo e o veículo de transmissão de idéia”. Nesse sentido, entender o termo “documento” e o seu amplo significado hoje, implica, antes de tudo, fazer uma busca nos postulados teóricos da história da Documentação para melhor compreender as transformações que ocorreram ao longo dos tempos.

No século XIX, o aumento no número de publicações contribuiu para a expansão da ideia de documento para muito além dos materiais bibliográficos (BUCKLAND, 1995, 1997). Ampliou-se o interesse para o desenvolvimento de métodos e técnicas de organização, análise de assunto, descrição e anotação bibliográfica que diferem marcadamente das práticas da biblioteca convencional. Essas novas técnicas passaram a ser chamadas de documentação (SHERA; CLEVELAND, 1977, p. 250), cuja sua principal função era a gestão e o controle da literatura, por meio de serviços de organização, representação, recuperação, reprodução e divulgação dos documentos.

Buckland (1997, p. 804) explica que inicialmente “[...] O termo tradicional para essa atividade foi ‘bibliografia’ [...]”, mas que a partir do século XX o uso da palavra documentação foi cada vez mais difundido na Europa para designar “[...] o conjunto de

técnicas necessárias para gerenciar essa explosão de documentos”. Depois de 1950, com a terminologia mais elaborada e direcionada para o contexto da Ciência da Informação, a palavra documentação surgiu em substituição ao antigo termo bibliografia para designar os serviços de armazenamento e recuperação da informação (BUCKLAND, 1997, p. 805).

Em 1934, com a publicação do *Traité de Documentation: le livre sur le livre*, a obra mais famosa de Paul Otlet, é que de fato ocorre a ideia de expandir o conceito de documento para além dos documentos textuais e dos registros gráficos, incluindo outros tipos de objetos.

[Documento] compreende não somente o livro, manuscrito ou impresso, mas revistas, jornais e reproduções gráficas de todas as espécies, desenhos, gravuras, cartas, esquemas, diagramas, fotografias, etc. A documentação, no sentido amplo do termo compreende: livros, elementos que servem para indicar ou reproduzir um pensamento, considerado sob qualquer forma (OTLET, 1934, p. 9).

Outro marco importante ressaltado por Shera e Cleveland (1977, p. 251) foi a contribuição dada pela francesa Suzanne Briet para a história da documentação, quando em 1951 publicou o manifesto *Qu'est-ce que la Documentation?*, no qual que indagava sobre a classificação de alguns objetos, se poderiam ou não ser considerados como documento.

[...] Uma estrela é um documento? Uma pedra jogada por uma torrente é um documento? Um animal vivo é um documento? Não. Mas, fotografias e catálogos de estrelas, os seixos num museu de mineralogia e catálogos de animais expostos num jardim zoológico, são documentos [...] (BRIET, 1951 apud SHERA; CLEVELAND, 1977, p. 251).

Assim, Briet propôs a seguinte definição para designar o documento, de forma mais abrangente:

O documento é todo objeto concreto ou sinal simbólico indexável, preservado ou registrado com a finalidade de representação, de reconstituição ou de prova de um fenômeno físico ou intelectual. [...] alguns documentos podem ser: as fotografias e os catálogos das estrelas, as pedras num museu de Mineralogia e os animais que são catalogados e mostrados num jardim Zoológico (BRIET, 1951, p.10).

Assim, diante do amplo significado de documento, enfatiza-se o periódico compreendido como documento, fonte de informação e parte integrante do processo de comunicação científica entre cientistas e o público (MEADOWS, 1999, p. 17; RIBEIRO; PINHEIRO; OLIVEIRA, 2007, p. 2).

Sobre este aspecto, considera-se que o processo de comunicação nos periódicos científicos pode operar tanto em suportes físicos de informação como em ambientes virtuais. Ou seja, sob quaisquer meios e canais de comunicação, do formal ao informal, do impresso ao eletrônico, desmistificando a ideia tradicional de documento, conforme observa Pinheiro:

A diversidade de conteúdos é a mesma de forma, isto é, o processo de comunicação se concretiza através dos mais diversos canais, tangíveis e intangíveis, orais, manuscritos, impressos ou eletrônicos. Objetos e fatos criados e promovidos pelo homem, por exemplo, contêm informação e fogem às categorias tradicionais de documento (PINHEIRO, 1998, p. 134).

A comunicação científica, tanto impressa como eletrônica, resulta de um processo evolutivo e de uma série de transformações de natureza histórica, conceitual e estrutural que ocorreram ao longo do tempo. Os produtos e serviços de informação não estão estritamente limitados às fontes convencionais de comunicação e divulgação de informações sob a forma impressa, tais como manuscritos, livros, periódicos e similares, mas também reúnem novos formatos disponibilizados *online* associados à coleta, ao processamento, ao arquivamento e a recuperação da informação de objetos digitais. Ou seja, aquilo que Otlet tão bem anunciava nos primórdios da documentação, sobre a expansão do conceito documento para muito além de seu suporte físico, suscitando outros formatos diferenciados de documento.

Mueller (2006, p. 27) explica que publicações científicas eletrônicas, incluindo os periódicos de acesso livre, trouxeram uma ruptura ao sistema tradicional de comunicação científica, oferecendo uma série de vantagens àqueles que necessitam fazer suas pesquisas de forma prática, ágil e gratuita.

[...] a grande maioria dos periódicos eletrônicos por assinatura, muito semelhantes, em aparência, ao modelo tradicional de periódico, com a importante diferença de serem acessíveis sem pagamento. Significam um ganho para o pesquisador em termos de acesso, conforto e presteza, mas não tanto em inovação, quando se considera a potencialidade do meio. Alguns títulos mantêm apenas a versão eletrônica para minimizar custos, e outros oferecem também uma versão impressa, paga (MUELLER, 2006, p. 32).

Lemos (2006), ao analisar as vantagens e desvantagens do periódico eletrônico, propõe a seguinte definição:

Um recurso eletrônico, com artigos completos, que pode incluir elementos de multimídia, disponível na internet, e que é publicado seqüencialmente, com uma designação numérica ou cronológica, e que pretende continuar indefinidamente. Pode ser a reprodução de uma revista impressa ou uma publicação exclusivamente em linha (LEMOS, 2006).

Fato é que na contemporaneidade, com a velocidade do fluxo de informação e o aumento expressivo da produção científica, o periódico eletrônico é um recurso facilitador do acesso livre à informação e ao conhecimento científico. O periódico eletrônico depositado numa *website* permite a desterritorialização do conteúdo, rompendo as distâncias, além de proporcionar a disseminação da informação e a educação científica dos usuários, conforme explicitam Oliveira *et al.*

O surgimento das publicações científicas eletrônicas representou importante elemento de dinamização na vida do pesquisador e da sociedade,

considerando que disponibilizam a informação de maneira rápida, com maior acessibilidade, em tempo real ou *online*, possibilitando maior visibilidade e custos reduzidos (OLIVEIRA *et al.*, 2010).

Na década de 60 houve uma preocupação com o crescimento exponencial da literatura técnico-científica. Saracevic (2009, p. 4) expõe que a partir da Segunda Guerra Mundial houve a inserção de uma variedade de tecnologias para controlar a explosão da informação, revelando que primeiramente foram introduzidos os cartões perfurados, o microfilme e, posteriormente, de forma mais massiva os computadores. Também esclarece as mudanças conceituais em torno da denominação “Recuperação da Informação” (RI), afirmando que com a influência da Ciência da Computação na CI, alguns nomes surgiram para nomear a RI como "literatura por máquina" e "organização mecânica do conhecimento", mas a expressão "**recuperação da informação**" foi a que realmente prevaleceu.

[...] Na verdade, a "*information retrieval*" (IR) foi inventada pelo matemático e físico Calvin N. Mooers (1919-1994), um pioneiro da computação e IR [...]. Ele postulou que Recuperação da Informação é [...] a descoberta ou processo de descoberta no que diz respeito às informações armazenadas [úteis ao usuário]. Recuperação da informação engloba os aspectos intelectuais da descrição da informação, sua especificação para busca e também qualquer sistema, técnica ou máquinas que são empregadas para realizar uma operação (MOOERS, 1957, p. 20-32 apud SARACEVIC, 2009, p. 5).

Quando Calvin Mooers cunhou o termo “recuperação da informação” contribuiu com ideias vigorosas para a CI, ao indicar que todos os aspectos relacionados ao processamento da informação devem ser considerados, desde a entrada até a saída de dados nos sistemas de recuperação da informação.

Atualmente, no ambiente da *Web Semântica*, novas mídias digitais estão sendo incorporadas às tecnologias de informação e comunicação (TICs), especialmente para usos sociais da ciência.

Até os dias de hoje, o sistema de comunicação de ciência depara-se com a relação crescimento da informação *versus* distribuição, sendo este um gargalo para as publicações impressas. Entretanto, as novas tecnologias de informação e comunicação eletrônica impactaram o sistema tradicional de comunicação da ciência (OLIVEIRA *et al.*, 2010).

3 CANAIS DE DIVULGAÇÃO DA CIÊNCIA COMO FONTES DE INFORMAÇÃO DIGITAIS

Na era digital, é sabido que a comunicação e divulgação do conhecimento científico têm sido realizadas intensamente em fontes eletrônicas ou a partir de tecnologias de informação e comunicação (TICs). Maciel ressalta que:

[...] A revolução científico-tecnológica da segunda metade do século XX, capitaneada pelas novas tecnologias de informação e comunicação (TICs),

atinge todas as esferas da vida humana de forma intensa e acelerada, principalmente a partir dos anos 70 do século passado [...] (MACIEL, 2005, p.107).

Nesse cenário, pode-se conjecturar que foi durante o século XX que a ciência definitivamente incorporou-se ao funcionamento cotidiano da sociedade. A ciência deixa de ser uma “instituição social heterodoxa” para desempenhar um papel estratégico como força produtiva e como mercadoria (ALBAGLI, 1996, p. 397). Com o desenvolvimento científico-tecnológico, aumenta a necessidade de melhor informar às pessoas a respeito da ciência e de seus impactos, favorecendo o aparecimento de iniciativas orientadas para políticas de divulgação científica.

Tornar amplamente acessível o conhecimento científico ao público é a principal função da divulgação científica, que na língua francesa é denominado de *vulgarisation scientifique* e na língua inglesa se diz *popular science (pop-science) ou Science popularization*. Desse modo, para melhor compreender o significado de divulgação científica e para estabelecer uma interface com os outros aspectos tratados neste texto é interessante partir das reflexões de Albagli, com base nos conceitos propostos por Bueno (1984), a saber:

Divulgação científica é um conceito mais restrito do que difusão científica e um conceito mais amplo do que comunicação científica. **Difusão científica** refere-se a "todo e qualquer processo usado para a comunicação da informação científica e tecnológica". Já **comunicação da ciência e tecnologia** significa "comunicação de informação científica e tecnológica, transcrita em códigos especializados, para um público seletivo formado de especialistas" (BUENO apud ALBAGLI, 1996, p. 399, grifo nosso).

A partir dessas distinções, entendemos que a difusão científica é o processo mais abrangente que envolve diferentes mecanismos de comunicação da informação científica e tecnológica, desde mensagens, serviços de comunicação oferecidos pelas unidades de informação e instituições de pesquisa, as comunicações orais e escritas apresentadas em congressos científicos e eventos similares, os programas de radiodifusão (BUENO, 1984; MASSARANI, 1998; MANSO, 2013, p. 16).

A comunicação científica refere-se à disseminação científica ou a transferência de informações entre os próprios cientistas, caracterizada por uma linguagem de especialidade codificada. Manso (2013, p. 16-17), com base em Bueno (1984), destaca que a comunicação científica pode ser categorizada em: “intrapares, quando praticada por especialistas de uma mesma área do conhecimento; e, extrapares, aquela direcionada para pesquisadores de outras áreas de conhecimento”. Já a divulgação científica é um conceito que se aproxima ao sentido de democratização e socialização do conhecimento científico, a partir do uso de uma linguagem acessível e inteligível direcionada ao público não especializado.

Desse modo, entende-se que a popularização da ciência ou divulgação científica trata-se da utilização de recursos de comunicação da informação científico-tecnológica destinada principalmente ao público leigo. Há ainda, o conceito de alfabetização científica (*scientific literacy*) que segundo Bauer (1994, p. 11) contempla os seguintes componentes culturais:

- Uma noção geral sobre determinados conceitos e temas substantivos da ciência;
- Uma noção sobre a natureza da atividade científica;
- Consciência do papel da ciência na sociedade e na cultura.

González de Gómez (2003) revela que esta concepção social da ciência remonta do final da década de 30, quando:

[...] a pressão estrutural das forças produtivas e a interpretação “universalizadora” do Estado, projetada pela revolução soviética, teve influência significativa sobre as comunidades científicas ocidentais, colocando em pauta a responsabilidade dos cientistas (GONZÁLEZ DE GOMEZ, 2003, p. 62).

Gradativamente, as mudanças significativas no segmento de popularização da ciência foram impulsionadas, principalmente, pelos seguintes motivos: a participação social no processo decisório relativo ao desenvolvimento científico e tecnológico; o crescimento significativo da produção científica recente; a complexidade da ciência e tecnologia e a necessidade de traduzi-la para não especialistas; o crescente *gap* de conhecimento científico e tecnológico; a perda de prestígio e de apoio público à ciência pela uma boa parcela da comunidade científica, em virtude da falta de compreensão social sobre essa atividade (ALBAGLI, 1996, p. 397-399).

Sabe-se que a produção do conhecimento atinge seu estágio mais completo quando ocorre à difusão por meio de canais de comunicação científica, condição fundamental para o avanço da ciência, geração de pesquisas e novos conhecimentos a serem divulgados ao grande público, constituindo-se numa verdadeira “educação científica” destinada ao público não especializado, conforme explica Maciel:

A educação científica tem, pelo menos, dois objetivos: trata-se da formação de novas gerações de cidadãos que tenham aquele conhecimento científico básico suficiente [...] e também de estimular e preparar o surgimento de novas levas de pesquisadores aptos a ampliar a autonomia do país em produção de ciência e de tecnologia. Mais uma vez, estamos nos referindo à inclusão social e ao desenvolvimento (MACIEL, 2005, p. 117).

No Brasil, segundo Pinheiro et. al. (2009, p. 3) “[...] as primeiras notícias sobre ciência divulgadas em jornais datam de 1808, ano da chegada da família real portuguesa e da fundação da Imprensa Régia [...]”. Posteriormente, o surgimento das universidades e a criação da Sociedade Brasileira para o Progresso da Ciência (SBPC), a prática da divulgação

científica foi mais intensificada. Para ilustrar, destacam-se o *Jornal Ciência Hoje* e a revista *Ciência Hoje* da SBPC, como bons exemplos de periódicos de divulgação científica.

No século XX, entre as instituições brasileiras dedicadas à divulgação científica, destaca-se o Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict/MCTI), com a criação do CanalCiência. Esse portal eletrônico disponibiliza atividades e serviços voltados ao ensino e à pesquisa científica (PINHEIRO; VALÉRIO; SILVA, 2009, p. 8-9). Essas iniciativas proporcionam a socialização do conhecimento e estimulam a criação de mais programas com fins democráticos no Brasil. Pinheiro (2009, p. 17) destaca a responsabilidade social do CanalCiência na educação científica em áreas do estado do Rio de Janeiro, ressaltando a participação em ações socioeducativas nas escolas:

O CanalCiência, desde a sua criação participa das Semanas Nacionais de Ciência e Tecnologia, promovidas pelo MCT, quando divulga o Portal e seus serviços, procurando levá-los até as Escolas, quando promove oficinas para professores e alunos do ensino médio e técnico, além de seguir a interiorização proposta pelo Ministério e se apresentado em cidades como Paracambi e na Baixada Fluminense (PINHEIRO, 2009, p. 17).

Do mesmo modo, González de Gómez explica que:

[...] nesta visão da ciência em sua função social, coloca-se a ênfase na comunicação científica, ficando o conhecimento associado ao conhecedor, de maneira indissociável. Acentua-se assim a potência de transformação da ciência enquanto formadora de competências, incorporadas logo nos sujeitos individuais ou coletivos (GONZÁLEZ DE GÓMEZ, 2003, p. 62).

4 AMAZONAS FAZ CIÊNCIA: PERIÓDICO DE DIVULGAÇÃO DA INFORMAÇÃO EM CIÊNCIA, TECNOLOGIA E INOVAÇÃO

Sob a responsabilidade do Departamento de Difusão do Conhecimento (Decon), unidade subordinada à Diretoria Técnico-Administrativa (Ditec), a Fapeam mantém periodicamente sua principal publicação, a *Amazonas Faz Ciência*, revista de divulgação da informação em CT&I destinada à socialização do conhecimento sobre pesquisas científicas do estado do Amazonas, em áreas estratégicas para o desenvolvimento da Amazônia. Segundo, o Art. 13, inciso VIII, da Lei Delegada N.º 116/2007, compete ao Decon (FAPEAM, 2013):

- Criação e gerenciamento de um sistema de informação em Ciência e Tecnologia no Estado do Amazonas;
- Assessoramento junto à imprensa;
- Coleta, tratamento e disseminação de informações em Ciência e Tecnologia;
- Coordenação de eventos relacionados CT&I e ao fomento da Fundação;
- Prestação de informações e assessoramento à Diretoria Técnico-Científica e ao Conselho Diretor nos assuntos inerentes à sua área de competência, especialmente

no oferecimento de subsídios à formulação de políticas públicas.

A revista é composta por uma diversidade de gêneros textuais do jornalismo científico, com textos informativos, de natureza descritiva, narrativa ou dissertativa, além de reunir algumas postagens de *email*, redes sociais (*facebook* e *twitter*), comentários, opiniões, iconográficos (imagens, fotografias, infográficos e *cartoons*), *hiperlinks* e *QR Code* de acesso a outras mídias digitais. Considerando somente os textos cujo conteúdo informativo é sobre a CT&I, os documentos concentram-se basicamente em dois (dois) tipos: artigos e entrevistas.

A revista *Amazonas Faz Ciência*, embora tenha sido idealizada em 2003, é um periódico de divulgação científica relativamente recente. Quando do ato da sua criação, em 2005, o título inicial da revista era *Amazonas Ciência* e, em 2007, passou a ser nomeada de *Amazonas Faz Ciência*, permanecendo inalterado até hoje. Nos primeiros anos, em função da fase de constituição e estruturação da revista, a publicação alternou entre uma periodicidade semestral, trimestral e quadrimestral. Desde 2010 passou a ser trimestral, alternando com tiragens de algumas edições especiais.

No que tange às características editoriais, percebe-se que a revista ainda está em franca expansão e em contínuo aprimoramento. Gradativamente vem incorporando os padrões de normalização relativos à editoração eletrônica. É composta por diferentes seções que apresentam artigos de divulgação sobre os mais variados assuntos que representam a identidade local e regional das comunidades amazônicas. Artigos estes que reúnem um conjunto de informações científicas divulgadas em linguagem acessível ao público não especializado.

Embora tenha passado por algumas mudanças, a estrutura da revista, praticamente ainda se mantém a mesma. Referimo-nos, de modo especial, as edições atuais da revista que apresentam as seguintes seções: a) essenciais: artigos de divulgação que abordam diversos assuntos de pesquisas científicas realizadas por pesquisadores bolsistas da agência de fomento, dentre os quais: Agronegócios, Antropologia, Arqueologia, Artes, Biocosméticos, Biodiversidade, Biofármacos, Biotecnologia, Cultura, Desenvolvimento Sustentável, Ecologia, Ecossistema, Ecoturismo, Educação, Geografia, História, Indústria, Inovação, Lingüística, Meio Ambiente, Manufaturas, Nanotecnologia, Recursos Energéticos, Saúde, Sociologia, Tecnologia, entre outras áreas; b) fixas: capa (chamadas principais e secundárias), espaço do leitor, canal ciência, descoberta, linguagem, multimídia, leitura acentuada, vida de cientista, a ciência responde, saúde, ciência na mão, entrevista e identidade.

A revista *Amazonas Faz Ciência* é uma publicação originada na forma impressa, também disponibilizada em meio eletrônico, sob o formato de texto PDF (*Portable Document Format*) que difere do formato de texto HTML (*Hyper Text Markup Language*) – o hipertexto – que é um formato de saída também adotado por algumas revistas *on-line* de acesso livre (*Open Access*). O PDF é um formato padrão para livros e revistas eletrônicas, por exemplo, os *eBooks* e congêneres. O HTML é um formato que segue uma linguagem de programação utilizada para criação de *home page*.

Ainda sobre os aspectos formais, nota-se que a revista está galgando um caminho que busca agregar inovações, com a inserção dos elementos interativos e materiais adicionais, totalmente personalizados e destinados à divulgação da ciência para o público infantil. Uma das vantagens deste tipo de revista de divulgação científica é a sua distribuição gratuita, permitindo que ela tenha um maior alcance na sociedade e popularização da ciência para o público não especializado.

Também merece destaque a publicação de algumas edições bilíngues da revista. Os números 24 e 28, respectivamente, com matérias intituladas “A luz que vem da Floresta” e “Fapeam 10 anos” foram disponibilizados à apreciação do público, tanto em língua portuguesa como na língua inglesa.

FIGURA 1 – Lançamento da *Amazonas Faz Ciência*



Fonte: Revista *Amazonas Faz Ciência*, nº 24, edição especial bilíngue na Rio+20, em 13 de junho de 2012. À esquerda, ministra do Meio Ambiente, Isabella Teixeira e à direita, diretora-presidenta da Fapeam, Profa. Dra. Maria Olívia de Albuquerque Ribeiro Simão. Fonte: Fapeam (2012).

Outro fato recente é o lançamento do protótipo da versão digital da revista para aparelhos de telefonia móvel (*mobile*) e os *tablets*, apresentado em 23 de junho de 2013, na 65ª SBPC, em Recife-PE.

Quanto à capa, observa-se que ao longo destes oito anos, houve mudanças gráficas, tanto no processo de diagramação como no *layout* da revista. Elementos como logotipo,

manchetes, vinhetas, verbetes, plano de fundo, fontes, cores, entre outros, relacionados ao estilo ou identidade gráfica da revista, paulatinamente foram sendo modificados. Acredita-se que, posteriormente, novos elementos devam ser acrescentados ou retirados, à medida que a revista atingir um grau de maturação no seu processo de editoração.



Figura 2 – Capa da revista *Amazonas Faz Ciência* da Fapeam, ano 2, n. 5, 2007.

Fonte: Portal Fapeam (2011).



Figura 3 – Capa da revista *Amazonas Faz Ciência* da Fapeam, ano 7, n. 22, out./dez. 2011.

Fonte: Portal Fapeam (2011).

Suplemento Infantil

A revista da Fapeam também conta com um material adicional destinado às crianças. Trata-se do Suplemento Infantil *Amazonas Faz Ciência Criança*, com textos de divulgação científica apropriados ao público infantil em idade escolar, apresentando curiosidades, descobertas, passatempos, depoimentos infantis e história em quadrinhos, sempre com o foco na ciência, no intuito de desmistificar, desde cedo, a ciência junto à sociedade, habituando a criança a se familiarizar com a leitura de textos que trazem assuntos científicos na linguagem mais adequada para esse tipo de público. Ao todo, foram publicados até o momento, nove edições desse suplemento, sendo que as quatro primeiras estavam na forma de encarte vinculado à revista principal e as cinco últimas edições, já sob a forma avulsa, editadas separadamente e com distribuição gratuita.

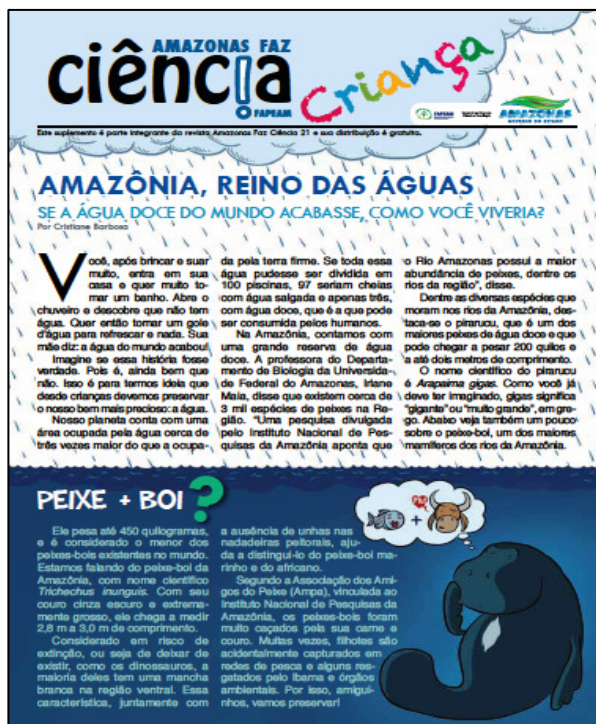


Figura 4 – Suplemento da revista *Amazonas Faz Ciência* da Fapeam, edição especial para crianças, 2011.

Fonte: Fapeam (2011).

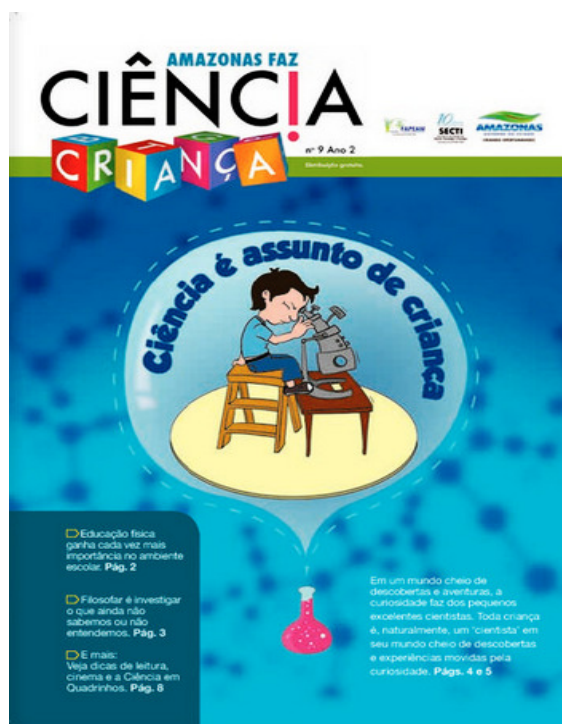


Figura 5 – Suplemento da revista *Amazonas Faz Ciência* da Fapeam, edição especial para crianças, 2013.

Fonte: Fapeam (2011).

5 CONSIDERAÇÕES FINAIS

Com base no estudo realizado, verificamos no periódico de divulgação científica *Amazonas Faz Ciência* sua adequação às políticas públicas da região e o papel desempenhado como fonte de divulgação de informações e socialização do conhecimento científico.

A partir de um levantamento exploratório da publicação, foi possível constatar que a revista de divulgação da informação em CT&I da Fapeam vem, ao longo dos anos, apresentando mudanças estruturais, tanto na forma como no conteúdo.

Pelos dados até aqui levantados referentes à revista de divulgação científica *Amazonas Faz Ciência*, indicam-se os seguintes resultados preliminares: a) a revista encontra-se no processo contínuo de aperfeiçoamento da sua estrutura, procurando adequar-se aos padrões de editoração de periódico de divulgação da informação em CT&I, oferecendo produtos inovadores para públicos específicos e agregando valor e qualidade à publicação; b) a revista *Amazonas Faz Ciência* é disponibilizada eletronicamente por meio de suporte digital; c) quanto ao estilo gráfico, a revista está passando por mudanças de editoração e, posteriormente, novos elementos poderão ser incorporados, à medida que ela for sendo consolidada.

Tendo em vista estas constatações iniciais quanto ao nível estrutural, também se pretende, a *posteriori*, analisar os elementos intrínsecos da revista, possibilitando a organização e a representação da informação em temáticas de domínios do conhecimento, além da recuperação da informação científica para estudos e pesquisas posteriores.

AGRADECIMENTOS

Agradecemos à Fundação de Amparo à Pesquisa do Estado do Amazonas (Fapeam) pela bolsa de pesquisa (Projeto original: A Revista *Amazonas Faz Ciência* da Fapeam e a Divulgação Científica em áreas estratégicas da Amazônia: uma proposta de análise bibliométrica, aprovado conforme Edital nº 020/2010, Conselho Diretor – Decisão nº 022/2011).

REFERÊNCIAS

- ALBAGLI, Sarita. Divulgação científica: científica para a cidadania?. **Ciência da Informação**, Brasília, v. 25, n. 3, p. 396-404, set./dez. 1996. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/view/465/424>>. Acesso em: out. 2010.
- BAUER, Henry Hermann. **Scientific literacy and the myth of the scientific method**. Chicago: University Illinois Press, 1994.
- BRAGA, Gilda Maria. Informação, ciência, política científica: o pensamento de Derek de Solla Price. **Ciência da Informação**, Brasília, v. 3, n. 2, p. 155-177, 1974.
- BRAGA, Gilda Maria; OBERHOFER, Alves. Diretrizes para a avaliação de periódicos científicos e técnicos brasileiros. **Revista Latinoamericana de Documentacion**, Brasília, v. 2, n. 1, p. 27-31, ene./jun. 1982.
- BRIET'S, Suzanne. **What is documentation?** Tradução de Ronald E. Day, Laurent Martinet e Hermina G. B. Anghelescu. Paris: Éditions Documentaries Industrielles et Techniques, 1951. Título original: Qu'est-ce que la documentation?. Disponível em: <<http://ella.slis.indiana.edu/~roday/what%20is%20documentation.pdf>>. Acesso: dez. 2009.
- BUCKLAND, Michael Keeble. What is a "documentation". **Journal of the American Society for Information Science - JASIS**, v. 48, n. 9, p. 804-809, 1997.
- BUCKLAND, Michael Keeble; LIU, Ziming. History of information science. **Annual Review of Information Science and Technology - ARIST**, v. 30, p. 385-416, 1995.
- BUENO, Wilson da Costa. **Jornalismo científico no Brasil: os compromissos de uma prática dependente**. 1984. 362 f. Tese (Doutorado em Ciências da Comunicação) – Escola de Comunicações e Artes, Universidade de São Paulo (USP), São Paulo, 1984.
- FERNANDEZ, Rosali Pacheco. Apresentação. In: LANGRIDGE, Derek. **Classificação: abordagem para estudantes de Biblioteconomia**. Rio de Janeiro: Interciência, 2006. p. 3-4.
- FUNDAÇÃO DE AMPARO À PESQUISA DO ESTADO DO AMAZONAS – FAPEAM. A Instituição. Manaus: a Fundação, Secretaria de Estado Ciência e Tecnologia, 2011-2013 Disponível em: <<http://www.fapeam.am.gov.br/pagina.php?cod=4>>. Acesso em: dez. 2013.

GONZÁLEZ DE GÓMEZ, Maria Nélide. As relações entre ciência, Estado e sociedade: um domínio de visibilidade para as questões da informação. **Ciência da Informação**, Brasília, v. 32, n. 1, p. 60-76, jan./abr. 2003. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/view/131/112>>. Acesso em: 14 out. 2010.

HERNANDEZ CAÑADAS, Patrícia Liset. **Os periódicos "Ciência Hoje" e "Ciência e Cultura" e a divulgação da ciência no Brasil**. 1987. Dissertação (Mestrado em Ciência da Informação) – Escola de Comunicação, Universidade Federal do Rio de Janeiro, Instituto Brasileiro de Informação em Ciência e Tecnologia, Rio de Janeiro, 1987. Orientadora: Profa. Dra. Heloísa Tardin Christóvão.

KRZYŻANOWSKI, Rosaly Fávero; FERREIRA, Maria Cecília Gonzaga. Avaliação de periódicos científicos e técnicos brasileiros. **Ciência da Informação**, Brasília, v. 27, n. 2, p. 165-175, maio/ago. 1998. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/view/357/318>>. Acesso em: set. 2011.

LANGRIDGE, Derek. **Classificação**: abordagem para estudantes de Biblioteconomia. Rio de Janeiro: Interciência, 2006. 120 p.

LEMONS, Antonio Agenor Briquet de. Periódicos eletrônicos: problema ou solução? DataGramZero – **Revista de Ciência da Informação**, Brasília, v. 7, n. 3, sem paginação, jun. 2006. Disponível em: <http://www.dgz.org.br/jun06/F_I_art.htm>. Acesso em: 30 set. 2014. Palestra pronunciada no X Encontro Nacional de Editores Científicos, São Pedro, SP, em 30 de novembro de 2005. Texto também disponível em: <<http://www.briquetdelemons.com.br/artigo07>>; <http://www.brapi.inf.br/_repositorio/2011/06/pdf_cd215b174d_0017331.pdf>. Acesso em: 30 set. 2014.

MACIEL, Maria Lucia. Estímulos e desestímulos à divulgação do conhecimento científico. In: Baumgarten, M. (Org.) **Conhecimentos e Redes**: sociedade, política e inovação. Porto Alegre: Ed. UFRGS, 2005. p. 107-117.

MANSO, Bruno Lara de Castro. **Divulgação científica e tecnológica**: interação entre agentes do processo. 2013. 104f. Dissertação (Mestrado em Ciência da Informação) – Escola de Comunicação, Universidade Federal do Rio de Janeiro, Instituto Brasileiro de Informação em Ciência e Tecnologia, Rio de Janeiro, 2013. Orientadora: Profa. Dra. Gilda Olinto. Co-orientador: Prof. Dr. Alfredo Tolmasquim.

MASSARANI, Luisa. **A divulgação científica no Rio de Janeiro**: algumas reflexões sobre a década de 20. 1998. 177 f. Dissertação (Mestrado em Ciência da Informação) – Escola de Comunicação, Universidade Federal do Rio de Janeiro, Instituto Brasileiro de Informação em Ciência e Tecnologia, Rio de Janeiro, 1998. Orientadora: Profa. Dra. Lena Vania Ribeiro Pinheiro. Co-orientador: Prof. Dr. Ildeu de Castro Moreira.

MEADOWS, Arthur Jack. **A comunicação científica**. Brasília: Briquet de Lemos/Livros, 1999. 268 p.

MIRANDA, Dely Bezerra de; PEREIRA, Maria de Nazaré Freitas. O periódico científico como veículo de comunicação: uma revisão de literatura. **Ciência da Informação**, Brasília, v. 25, n.3, p. 375-382, set./dez. 2005. Número especial: IBICT 50 anos. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/view/462/421>>. Acesso em: 14 out. 2010.

MUELLER, Suzana Pinheiro Machado. A comunicação científica e o movimento de acesso livre ao conhecimento. **Ciência da Informação**, Brasília, v. 35, n. 2, p. 27-38, 2006. Disponível em: <http://www.dgzero.org/fev05/Art_02.htm>. Acesso em: mar. 2013.

OLIVEIRA, Eloísa da Conceição Príncipe; VALERIO, Palmira Moriconi; CHALHUB, Tânia; PINHEIRO, Lena Vania Ribeiro. Recursos eletrônicos de informação e comunicação em periódicos científicos latino-americanos em Ciências Agrárias. In: JORNADAS LATINOAMERICANA DE ESTUDIOS SOCIALES DE LA CIENCIA Y LA TECNOLOGIA – CIÊNCIA Y TECNOLOGIA PARA LA INCLUSIÓN SOCIAL EM AMÉRICA LATIN, 8., 2010, Buenos Aires. **Anais...** No prelo.

OTLET, Paul. *Traité de documentation: le livre sur le livre, théorie et pratique*. Bruxelles, Belgium: Ed. Mundaneum, 1934.

PEREIRA, Maria de Nazaré Freitas. Prefácio que esclarece o leitor a propósito do sonho de Otlet: aventura em tecnologia da informação e comunicação. In: PEREIRA, Maria de Nazaré Freitas; PINHEIRO, Lena Vania Ribeiro. **O sonho de Otlet: aventura em tecnologia da informação e comunicação**. Rio de Janeiro: IBICT, 2000, p. vii-xxiv.

PINHEIRO, Lena Vania Ribeiro. Campo interdisciplinar da Ciência da Informação: fronteiras remotas e recentes. **Investigación Bibliotecológica**, México, v. 12, n. 25, p. 132-163, 1998. Disponível em: <<http://biblioteca.ibict.br/phlB/anexos/Mexicolena.pdf>>. Publicado também no Brasil em coletânea do IBICT.

PINHEIRO, Lena Vania Ribeiro; BRASCHER, Marisa; BURNIER, Sonia. Ciência da Informação: 32 anos (1972-2004) no caminho da história e horizontes de um periódico científico brasileiro. **Ciência da Informação**, Brasília, v. 34, n. 3, p. 25-77, set./dez. 2005. Número especial: IBICT 50 anos. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/view/816/657>>. Acesso em: 14 out. 2010.

PINHEIRO, Lena Vania Ribeiro; SILVA, Márcia Rocha da; SOUZA, Sonia Burnier de; BARROS, Flavia Rubenia da Silva; GUERRA, Claudia Bucceroni. Experiência inovadora do CanalCiência como instrumento pedagógico para aproximar ciência e sociedade, conhecimento e informação. **DatagramaZero**, Rio de Janeiro, v. 10, n. 4, set./out. 2009. Disponível em: <http://www.dgz.org.br/out09/Art_02.htm>. Acesso em: 15 out. 2010.

PINHEIRO, Lena Vania Ribeiro; VALERIO, Palmira M.; SILVA, Márcia R. Marcos históricos e políticos da divulgação científica no Brasil. In: BRAGA, Gilda Maria; PINHEIRO, Lena Vania Ribeiro (Org.). **Desafios do impresso ao digital: questões contemporâneas de informação e conhecimento**. Brasília: IBICT; Unesco, 2009. p. 259-289.

RIBEIRO, Carolina Krause; PINHEIRO, Lena Vania Ribeiro; OLIVEIRA, Eloísa da Conceição Príncipe de. Construção de um modelo-síntese para análise de periódicos científicos. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 8., 2007, Salvador. **Anais...** Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação (ANCIB), 2007. Disponível em: <<http://www.ancib.org.br/media/dissertacao/GT7--156.pdf>>. Acesso em: dez. 2011.

SARACEVIC, Tefko. Information science. In: BATES, Marcia J.; MAACK, Mary Niles (Ed.). **Encyclopedia of Library and Information Science**. New York: Taylor & Francis, 2009.

SARACEVIC, Tefko. Relevance: a review of and a framework for the thinking on the notion in Information Science. In: BRAGA, Gilda; PINHEIRO, Lena Vania Ribeiro (Org.). **Desafios do impresso ao digital: questões**. Brasília: Ibict: Unesco, 2009. p. 15-30.

SHERA, Jesse Hauk; CLEVELAND, Donald B. History and foundations of Information Science. **Annual Review of Information Science and Technology - ARIST**, v. 12, p. 249-275, 1977.

SOUZA, Rosali Fernandez de. Organização e representação de áreas do conhecimento em Ciência e Tecnologia: princípios de agregação em grandes áreas segundo diferentes contextos de produção e uso de informação. **Enc. Bibli. R. Eletr. Bibliotecon. Ci. Inf.**, Florianópolis, número especial, p. 27-41, jan./jun. 2006. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2006v11nesp1p27/384>>. Acesso em: 25 abr. 2013.

STUMPF, Ida Regina Chitto. Passado e futuro das revistas científicas. **Ciência da Informação**, Brasília, v. 25, n. 3, p. 383-386, set./dez. 1996. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/view/463/422>>. Acesso em: 15 out. 2010.

INSTITUCIONALIZAÇÃO CIENTÍFICA DA ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO NA CIÊNCIA DA INFORMAÇÃO NO BRASIL

Gracy Kelli Martins
João Batista Ernesto de Moraes

Resumo: A organização e representação do conhecimento (ORC) possui raízes epistemológicas em tempos longínquos. Está entre as preocupações de âmbito humano, desde as antigas civilizações. Seus processos foram tomados como arte, técnica até transformar-se em disciplina científica. Hoje, encontra-se no cerne da Ciência da Informação (CI) e direciona grande parte de suas pesquisas teóricas e instrumentais com sua institucionalização científica estabelecida. Na CI brasileira é nominada pela expressão Organização e Representação do Conhecimento, tendo em vista que mesmo sendo considerada uma subdisciplina, ela também se constitui como um campo independente de investigação que permeia outras áreas científicas. Como subdisciplina da CI, no país, apresentou-se primeiramente através de campos de estudo da biblioteconomia e documentação até constituir-se como área de pesquisa científica no interior dos programas de pós-graduação. Este trabalho expõe o percurso histórico de suas práticas e teorias e seu desenvolvimento no Brasil, no âmbito da CI, evidenciando sua institucionalização científica pautada na formação especializada e no desenvolvimento de um crescente número pesquisa na área.

Palavras-chave: Ciência da Informação. Institucionalização Científica. Organização e Representação do Conhecimento.

Abstract: The knowledge organization and representation (KOR) has epistemological roots in ancient times. It is among the concerns of human affairs, since ancient civilizations. Its processes were taken as art, technique until transform itself into a scientific discipline. Today, it lies at the core of information science (IS) and directs much of its theoretical and instrumental research with its scientific institutionalization established. In Brazilian IS it is nominated by the expression Knowledge Organisation and Representation, considering that even being considered a subdiscipline, it also represents an independent research field that permeates other scientific areas. As a subdiscipline of IS in the country, was presented primarily through field study of librarianship and documentation until constitute itself as an area of scientific research within the postgraduate programs. This paper presents the historical background of their practices and theories and its development in Brazil, within the IS, demonstrating its scientific institutionalization guided by specialized training and the development of an increasing number of research in the area.

Keywords: Information Science. Knowledge Organisation and Representation. Scientific institutionalization.

1 INTRODUÇÃO

A organização e representação do conhecimento (ORC) possui raízes epistemológicas em tempos longínquos. Sempre esteve entre as preocupações de âmbito humano, desde as civilizações anteriores a Cristo, além de permear os mitos, as investigações metafísicas, matemáticas, biológicas e históricas. Foi aprofundada por filósofos, discutida por documentalistas e bibliotecários, ordenou acervos em mosteiros, gabinetes de curiosidade e arquivos contábeis na Mesopotâmia. O registro de maior relevância lendária, referente aos

processos de organização documental, encontra-se expresso na narrativa da maior Biblioteca da história, a Biblioteca de Alexandria (ORTEGA, 2004).

Um dos fatores mais marcantes entre as civilizações ocidentais foi o desenvolvimento da escrita. Esta tecnologia permitiu não só uma evolução cultural como também, o aprimoramento dos registros do conhecimento. Toda evolução histórica ligada à escrita impulsionou avanços tecnológicos na produção de livros e documentos, exigindo a “criação de instrumentos capazes de classificar, identificar e hierarquizar o conhecimento disponível” (PINHO, p. 24, 2009).

Tais avanços instituíram impactos culturais, tecnológicos, científicos e econômicos, evidenciando cada vez mais a necessidade de representar e organizar através da preservação, codificação, decodificação, compartilhamento e uso dos registros do conhecimento. Dessa forma, a ordenação física e as investigações filosóficas, sobrepostas nos processos de organização documental culminaram em um campo de estudos que contemplou a investigação científica em um âmbito reconhecidamente interdisciplinar, e prestou-se a resolver os problemas de organização e acesso ao conhecimento (SOUZA, 1998; ZANDONADE, 2003; PINHO, 2011). Entretanto, qualquer tentativa de reconstrução histórica referentes à ORC não poderá furtar-se de sua essência epistemológica junto à Biblioteconomia e à Documentação, no contexto brasileiro em específico.

Os registros históricos evidenciam que os estudos em organização e representação do conhecimento caracterizaram-se como atividade primeira da Biblioteconomia e Documentação, e conseqüentemente da CI, desenvolvidos na EUA e Europa, respectivamente. Com o despontar do processo acelerado de publicações, surgem os problemas de catalogação e classificação de periódicos científicos e o aprimoramento dos processos de organização e representação para além dos espaços das bibliotecas. De acordo com Zandonade (2003) em meados do século XIX, essa produção científica já acumulava em torno de dez mil títulos, possibilitando que a Documentação, ao final desse mesmo século apresentasse suas grandes contribuições para a solução dos problemas relativos ao documento e não mais só ao livro. No fim do século XIX Biblioteconomia e Documentação historiam um desenvolvimento imbricado, uma vez que: [...] surgiram em consequência das mesmas necessidades, empregavam processos e instrumentos comuns (como as fichas de 7,5 por 12,5 cm e a Classificação Decimal de Dewey-CDD), tinham objetivos quase idênticos e em muitos casos deviam seu progresso aos mesmos homens (ORTEGA, 2004, p. 4).

Fixam-se nesse contexto as múltiplas direções de estudo da ORC¹⁷², permitindo que sua investigação tome por abordagem a teoria do conhecimento “com o estabelecimento de divisões de tipologias de conhecimento” e o “enfoque nas formas idiossincráticas de definição de significado social a termos linguísticos” até a “elaboração de sistemas de organização voltados a domínios específicos de conhecimento” (SEMIDÃO; ALMEIDA; MOREIRA, 2013, p. 591).

2 INSTITUCIONALIZAÇÃO CIENTÍFICA DA ORC

Os sistemas de comunicação científica se expandiram consideravelmente tornando-se um instrumento de promoção individual e fortalecimento do ciclo de organização e difusão do conhecimento, onde a institucionalização científica de uma área é evidenciada como fator determinante para a produção de novas e interdisciplinares pesquisas (BUFREN, et al, 2007). Conduzida por essa dinâmica evolutiva da ciência, a Ciência da Informação institucionaliza-se a partir do desenvolvimento social e tecnológico relacionados diretamente ao fenômeno da informação. Seu estabelecimento, enquanto campo científico, refere-se a uma ciência que surge como resposta para as transformações tecnológicas e a necessidade social de organizar, armazenar e recuperar informações (PEDROSO IZQUIERDO, 2004).

Whitley (1974) expõe que para avaliar uma área científica é necessário que se meça sua institucionalização através de: processos desenvolvidos; grau de coerência; clareza do arcabouço teórico e metodológico; estrutura conceitual; organização de ações mediante divulgação de resultados; formação especializada e a articulação entre seus pesquisadores com espaços específicos de discussão. A análise da institucionalização científica de uma área permitirá identificar com clareza as sociedades científicas e quem são seus pesquisadores, quais reuniões profissionais divulgarão os resultados de suas investigações e quais periódicos serão responsáveis pela sua comunicação científica.

Para o autor, a institucionalização deve ser pautada por duas dimensões: a institucionalização cognitiva e a institucionalização social. Nessa perspectiva, o conceito de institucionalização refere-se a padrões e significados e não há separação entre a

¹⁷² Adotar-se-á OC quando mencionada a grande área de múltiplos estudos, que permeia todas as ciências, e ORC quando referir-se especificamente à área de investigação dentro da Ciência da Informação. A opção por “Organização e Representação do Conhecimento”, no decorrer do texto, orienta-se por esta ser a nomenclatura adotada para definir o Grupo de Trabalho (GT2) no interior da Associação Nacional de Pesquisa e Pós-graduação em Ciência da Informação - ANCIB, no Brasil, que trabalha com a especialidade em questão. Também é assim reconhecida no âmbito científico da *International Society for Knowledge Organization* – ISKO, tanto nacional, como internacionalmente.

institucionalização cognitiva e social, já que ambas encontram-se na esfera social do fazer da ciência. “O nível de coerência e organização de ações e percepções e a extensão das articulações à aderência das ideias constituem o nível de institucionalização” (WHITLEY, 1974, p. 71, tradução nossa).

Em específico, a institucionalização cognitiva, a partir de seu conjunto de situações-problema, busca identificar se há, dentro da área analisada, semelhanças cognitivas e consenso quanto às investigações de seus fenômenos, sem a necessidade de dispor de definições idênticas. Nessa avaliação, quanto maior o consenso na definição do objeto/fenômeno, incluindo seu arcabouço conceitual, e aplicação de técnicas apropriadas e resultados significativos, maior será o nível de institucionalização cognitiva da área científica. A institucionalização social evidenciará a formação especializada, as redes de relacionamento, a divulgação de resultados de pesquisa e a cooperação entre pesquisadores e especialidades. Não é incomum que haja variação de níveis entre a institucionalização cognitiva e social, no entanto, uma sempre será dependente da outra na garantia da institucionalização científica (WHITLEY, 1974).

Hjørland (2002) também toma como base as considerações postas por Whitley (2000), reconhecendo que a avaliação de uma determinada área sempre precisará adotar a perspectiva organizacional, tendo em vista que possui uma dependência mútua entre as estruturas de comunicação científica e o âmbito social e sugere que esta perspectiva também seja útil à CI, para que os índices avaliados identifiquem o processo de estruturação da área científica para além de dados puramente estatísticos (HJØRLAND, 2002).

Esta institucionalização¹⁷³ Bourdieu (2004) define como estruturação do campo científico, e está alicerçada pelos indivíduos e pelas instituições responsáveis, através do volume do capital científico que determina a estrutura do campo. Este capital, atendendo às questões específicas de determinado campo, compõe não somente suas pesquisas científicas, através de descobertas e consolidação das já existentes, como também, solidifica as relações políticas e sociais dentro e entre as instituições, garantido a solidez através de atos de conhecimento e reconhecimento no interior do campo, que são os processos institucionalizados de avaliação pelos pares (BOURDIEU, 2004).

¹⁷³ A definição de campo científico, proposta por Pierre Bourdieu não se difere do autor inglês Richard Whitley, na conceituação de área científica. A compreensão para a utilização de ambos se dá por reconhecer que Bourdieu parte de uma análise sócio filosófica, essencial para o aprofundamento teórico da discussão, à medida que, Whitley propõe, de forma objetiva, critérios para a investigação e consolidação dos dados necessários para a análise da institucionalização científica.

A institucionalização “formal moderna” da Biblioteconomia, como responsável principal pela “gerência da informação, ou mais precisamente, a de organizadora e administradora de coleções bibliográficas” (SOUZA, 1998, p.9), foi reivindicada e reconhecida a partir da criação da *American Library Association*, em 1876 e no ano de 1887 com o início do *Course of Library Economy, no Columbia College*, proposto por Melvil Dewey. Entretanto, seu marco como uma ciência da Biblioteconomia estabeleceu-se com a Escola Graduada de Biblioteconomia (Graduate Library School – GLS) na Universidade de Chicago, em 1926. Depois da Segunda Guerra Mundial, “as técnicas de informação tornaram-se generalizadamente disponíveis para a profissão da Biblioteconomia” transformando a “recuperação da informação” a principal corrente de investigação e prática nos Estados Unidos. (ZONDONADE, 2009)

Henry Evelyn Bliss foi responsável pela publicação de obras que incidiram em grandes contribuições para área de organização e representação do conhecimento. Foi também o autor que utilizou formalmente e, de maneira notória, a expressão organização do conhecimento, dentro da área de Biblioteconomia e, posteriormente, da Ciência da Informação em suas publicações “*The organization of knowledge and the system of sciences*”, em 1929, e “*The organization of knowledge in libraries*”, em 1933 (GUIMARÃES, PINHO, FERREIRA, 2012). Murguia e Sales (2013, p. 447), pontuam que além de Bliss, o termo foi utilizado também por Paul Otlet, em um artigo de 1903, intitulado “A ciência da Bibliografia e da Documentação”.

As exigências oriundas do desenvolvimento científico e tecnológico impulsionaram comunidades científicas que encaminharam a Biblioteconomia para um núcleo de investigação que não tinha mais o livro como foco e sim, a informação. Assim, por volta dos anos de 1950 nasce uma Ciência da Informação que buscou interdisciplinaridade com áreas também nascentes dessa época, como “a ciência da computação, cibernética, teoria geral de sistemas, teoria da informação, lógica formal, etc.” (ZANDONADE, 2003, p.36).

Pioneiro nos estudos da ciência da Biblioteconomia, Documentação e Ciência da Informação (nessa ordem, respeitando a evolução científica de cada época), Shera desenvolveu significativas contribuições para uma institucionalização da Biblioteconomia e Ciência da Informação como disciplina científica, cujo objetivo ele entendia ser a “organização dos registros gráficos do conhecimento para o seu melhor uso possível” (EGAN; SHERA, 1951, *apud* ZANDONADE, 2003, p. 51).

Com suas bases estabelecidas, a Biblioteconomia assumiu o núcleo de trabalho que a definia: “a) organização prática do conhecimento para fins de fornecer respostas aos

utilizadores de informação; e b) organização e administração de coleções bibliográficas”. Diante de tais processos tornou-se necessário que este profissional adotasse critérios e fundamentasse uma teoria para organizar tais conhecimentos e administrar as coleções a fim de tornar “acháveis os pacotes de conhecimento” (SOUZA, 1998, p.9).

Prosseguindo nos registros históricos de Souza (1998, p.9-10), o autor pontua que tal teoria não se originou do “auxílio imediato das sistematizações filosóficas ou científicas, de organização do conhecimento, ainda que nelas inspirada”. Das experimentações empíricas na relação sociedade/coleções de livros, originaram-se instrumentos que “tinham como finalidade o armazenamento dos livros por sua descrição e os que tinham por fim último a recuperação dos livros segundo seu conteúdo”. Processos utilizados até os dias atuais, a catalogação e a classificação se corporificam em diretrizes para a organização de documentos, mediante a descrição formal destes e a recuperação segundo seu conteúdo, respectivamente.

Tais Regras de catalogação e Tabelas de Classificação, elaboradas pelos peritos da organização prática do conhecimento, permitem compor e dispor uma coleção documental. Elas dão a base para um certo tipo de organização do conhecimento de caráter não-filosófico e não científico. Seu objetivo é justificadamente pragmático, conforme foi feito pelo bibliotecário norte-americano Melvil Dewey, por exemplo. É essa organização pragmática do conhecimento que vem a se transformar em uma teoria da classificação bibliográfica. Ela viabiliza também a constituição de uma teoria da Organização e Administração de bibliotecas e dá origem a uma nova “ciência”, a Biblioteconomia, que veio a ser convertida em sustentação para uma profissão formal, a de bibliotecário (SOUZA, 1998, p. 11).

Muitas nomenclaturas foram adotadas para definir o exercício da organização e representação do conhecimento, em especial na área de Biblioteconomia e Documentação, no qual é possível reconhecer a Biblioteconomia como uma das forças impulsionadoras da organização e representação do conhecimento na sociedade moderna, em transição para a pós-modernidade.

Um exemplo claro, presente na literatura, refere-se ao termo Bibliografia, bastante adotado a partir de meados do século XIX e por todo século XX e empregado como um dos instrumentos primordiais do controle bibliográfico. Este último, utilizado por Shera e Egan definia os processos de Documentação, ainda nos anos de 1949, quando, acompanhando as mudanças, inclusive profissionais da época, passam a tomar em uso o termo “organização bibliográfica”, observando que ambos permanecem até hoje na literatura e designam segundo seus criadores: uma “teoria da armazenagem e recuperação da informação”, considerada por Shera como sendo o cerne da Ciência da Informação” (ZANDONADE, 2003, P.52, grifo do autor).

Ao longo dos tempos, muitas áreas intentaram estabelecer regras e procedimentos acerca da organização do conhecimento, enquanto disciplina. Inicialmente essas investigações eram da alçada da filosofia, entendendo neste aspecto o conhecimento como materialização das compreensões humanas. O registro e a materialização desse conhecimento em um suporte físico passaram a ser o objeto das disciplinas responsáveis pelo armazenamento, guarda e recuperação, do que Souza (1998) chama de “pacotes de conhecimento”.

Em todo seu percurso evolutivo essas práticas, científicas ou pragmáticas, receberam distintas definições até firmarem-se como Organização e Representação do Conhecimento, como é adotado atualmente, destacando as mais evidenciadas na literatura: Documentação (OTLET, 1934); Controle Bibliográfico (EGAN; SHERA, 1949); Organização Bibliográfica (EGAN; SHERA, 1951); Organização da informação (*American Documentation*, 1950, apud ZANDONADE, 2003); Ciência da Documentação (LOPEZ YEPES, 1978); Organização do Conhecimento (DAHLBERG, 2006). Dessa forma, a organização dos registros do conhecimento conquista espaço, não só como uma atividade pragmática, mas expande-se em um campo de estudos teóricos e aplicados e suas contribuições dão origem a uma disciplina tornando assim, mais acessível o uso e a apropriação do conhecimento produzido.

Como já mencionado, as teorias em ORC não se originaram de imediato das sistematizações filosóficas ou científicas. Suas técnicas nascem na diversidade da organização de documentos e se dividem em natureza descritiva: “organização dos documentos enquanto produtos fisicamente formatados” e, em natureza temática: “organização temática dos documentos a partir do conhecimento neles contido” (GUIMARÃES, 2003; SOUZA, 1998, p. 55). Nessa perspectiva, de abordagens em dois diferentes níveis, Guimarães (2008, p.78) ressalta que o Tratamento Temático da Informação – TTI –, junto à Ciência da Informação e a Biblioteconomia:

Ocupa (como se pode observar tanto na literatura quanto nas distintas práticas profissionais) um espaço nuclear, visto revelar a mediação entre a produção e o uso da informação, entre elas tecendo a mais sólida ponte: a que dá acesso ao conteúdo informacional.

No que diz respeito ao seu tratamento temático, muitos sistemas de organização do conhecimento foram estudados e aplicados e se delinearam, tanto de forma pragmática, como científica, em três correntes teóricas: catalogação de assuntos, indexação e análise documental (GUIMARÃES, 2008). Tais correntes foram estruturadas, de acordo com Guimarães, Moraes e Guarido, (2008, p. 5), a partir da sistematização de conteúdos em organização e

representação e se apresentam da seguinte forma: a) **processos**: compreendem *a análise*¹⁷⁴, *a condensação*¹⁷⁵ e *a representação*¹⁷⁶; b) **produtos**: a transformação dos documentos em produtos que facilitam a consulta aos originais, em áreas especializadas do conhecimento, e; c) **instrumentos**: ferramentas de linguagem documental para a representação padronizada do conteúdo temático de documentos, fazendo uso acentuado das tecnologias de informação e comunicação.

A consolidação acadêmica da Biblioteconomia e, posteriormente, da Ciência da Informação e todo aparato tecnológico que desponta a partir dos anos 1950, desencadearam experiências de tratamento automatizado da informação (GUIMARÃES, 2003; 2008). Assim, a organização e representação do conhecimento ganhou “espaço não apenas como necessidade pragmática para o universo documental, mas como campo de reflexão e produção teórica” (PINHO, 2009, p. 30). Para além de tais experiências, surge a preocupação com os estudos de bases científicas para “o fazer do TTI, no intuito de ir além de técnicas prescritivas para buscar a construção de metodologias defensáveis para o desenvolvimento dos procedimentos da área” (GUIMARÃES, 2008, p 80).

Partindo de técnicas prescritivas e originando teorias, as contribuições advindas dos estudos de Ranganathan, em especial das descrições acerca do desenvolvimento da *Colon Classification*, influenciaram, ainda no ano de 1952 a fundação da *Classification Research Group*, na Inglaterra. O grupo tinha por finalidade aperfeiçoar e melhorar as ferramentas existentes voltadas para a classificação da informação técnico-científica e estabelecer tais estudos enquanto campo de estudo específico. Entretanto, findou seus trabalhos em 1968, e contou com grandes nomes da literatura especializada em ORC, como Derek Austin, Eric Coates, Jason Farradane, Robert Fairthorne, Douglas Foskett, Barbara Kyle, Derek Langridge, Jack Mills, Bernard Palmer, Jack Wells, e Brian Campbell Vickery.

A institucionalização científica da ORC tem como um dos principais marcos a *Society for Classification*, fundada em 1977 por Ingetraut Dahlberg, que abrangia os estudos e aplicações voltados para “a classificação, por meio de tesouros, taxonomias para ordenar objetos e taxonomia numérica, usando métodos formais e matemáticos para o reconhecimento de classes de objetos de acordo com suas características” (PINHO, 2009, p. 30). A

¹⁷⁴ Parte da leitura de documentos, buscando as partes de maior relevância temática, exigindo assim um conjunto de estratégias metacognitivas;

¹⁷⁵ Trabalha no sentido de síntese do documento por termos de recuperação;

¹⁷⁶ Traduz o conteúdo temático do documento a uma linguagem especializada, fazendo uso da indexação por índices, vocabulários controlados, tesouros, entre outros.

composição do grupo contava com matemáticos e estatísticos que focavam as investigações em métodos de análise de dados. Não sendo esse o único interesse de Dahlberg, a pesquisadora, após 12 anos de contribuições, deixou a *Society for Classification* para dar início a um novo e decisivo projeto para a área de ORC: a *International Society for Knowledge Organization*, em 1989.

De acordo com Guimarães (2008); Pinho (2009) e Ribeiro e Cerveira (2013), é através da *International Society for Knowledge Organization* – ISKO – que o universo investigativo e a produção teórica em ORC tem se configurado como campo científico, propiciando principalmente a interlocução científica internacional, reconhecendo a ORC como um dos marcos conceituais da CI, inclusive evidenciando historicamente seus aspectos teóricos, conceituais e instrumentais que envolvem a organização do conhecimento no cerne desta Ciência.

A ISKO foi fundada no ano de 1989, aos 22 dias do mês de julho, em Frankfurt, na Alemanha. Sua fundadora, Ingetraut Dahlberg, decidiu criar uma sociedade voltada para a organização e representação do conhecimento, no qual a ORC teria uma abordagem mais ampla, investigando o conhecimento a partir de um melhor entendimento de suas características organizacionais, descritivas, representativas e sua aplicação a partir do acesso e uso desse conhecimento (PINHO, 2009).

Dahlberg (2006) elucida que a escolha pelo nome da sociedade, adotou como referencial as obras de Bliss, que tomaram por uso a expressão (*Knowledge Organization*) em suas publicações, nos anos de 1929 e 1933, como já mencionados. Para a pesquisadora, dentro do escopo em que se firmaria a ISKO, a organização do conhecimento compreenderia “os objetos e atividades da teoria do conceito, classificação e indexação, e representação do conhecimento” (DAHLBERG, 2006, p. 12, tradução nossa). Esclarece que, por representação do conhecimento a ISKO entendia não só a estrutura lógica das representações conceituais, como também todos os problemas relativos aos conceitos e termos adequados, uma vez que as questões de terminologia também tinham que ser consideradas (DAHLBERG, 2006).

Ao longo dos seus 25 anos de existência, a ISKO ramificou-se em capítulos nacionais e regionais e, possui membros em mais de cinquenta países, das mais diversas especialidades (Ciência da Informação, Filosofia, Linguística, Ciência da Computação, entre outras). Sua atuação tem como foco os problemas relacionados à organização conceitual, para além da necessidade pragmática do universo documental, constituindo-se como um amplo e representativo fórum científico internacional.

Atualmente, conta com a realização de suas reuniões nos seguintes países: Brasil, Canadá, EUA, China, França, Alemanha, Índia, Irã, Itália, Maghreb (Argélia; Tunísia; Marrocos), Polônia, Espanha, Reino Unido, além de correspondentes na Austrália, Hungria, Geórgia, Norte da Europa, Romênia, Rússia e Eslováquia¹⁷⁷. Seus eventos regionais e nacionais são bienais, e congregam a discussão de temas específicos, bem como propicia a reunião entre seus membros, oferecendo associações individuais e institucionais.

Nesse contexto, a literatura evidencia que os principais autores da área de ORC estão vinculados não só aos conglomerados científicos que constituem a especialidade em Organização do conhecimento – apontando um alto nível de institucionalização social por apresentar áreas de pesquisa claramente delineadas e em diversificadas cooperações – como também, dentro desta, um desenvolvimento cognitivo demarcado a partir de investigações mediante aos modelos dominantes e constante investigação por novos modelos de análise, para maior compreensão de fenômenos definidos no seu interior.

Tais considerações se pautam no que Souza (1998, p.46) já apontava na década de 1990, quando expunha que se enfrentaria uma crise pelo “próprio desenvolvimento da ciência e da técnica, na medida em que seus limites e referenciais epistemológicos não se sustentam diante dos novos modos de produção”. E nesse prenúncio, os mesmos fatores materiais ligados à ORC assumiriam nova dinâmica e provocariam outras novas ordens que balizariam diferentes formas de organização documental, modificando dessa forma certas técnicas tradicionais, em especial da Biblioteconomia (SOUZA, 1998).

3 ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO NA CIÊNCIA DA INFORMAÇÃO BRASILEIRA

Não é possível abordar qualquer tipo de problema relativo à institucionalização científica de uma área sem antes deitar o olhar sobre sua epistemologia, em busca da compreensão de sua etimologia e história. No Brasil a ORC, enquanto campo científico é introduzida através dos currículos de Biblioteconomia. Conquanto alguns pesquisadores da CI procurem desvinculá-la da Biblioteconomia, é impossível negar seus vínculos históricos e em especial no país, onde a implantação da CI ocorre nas escolas de ensino superior em Biblioteconomia, já existentes, no qual inicialmente foram implantados os programas de pós-graduação nominados como “Mestrado em Biblioteconomia”, ao exemplo da UFMG e UFPB, sendo substituídos apenas por volta dos anos 90 (SOUZA, 2012).

¹⁷⁷ <http://www.isko.org/events.html>

Embora a pesquisa substantiva e sistemática em CI no país tenha tido início nos programas de pós-graduação da área, ligadas aos cursos de Biblioteconomia, é inegável que o desenvolvimento da CI ocorre também no âmbito da graduação (ARBOIT; BUFREM; KOBASHI, 2011, p. 145).

O primeiro curso surgiu a partir do Decreto 8.835 de 11 de Julho de 1911, na Biblioteca Nacional, com influência da escola francesa *École de Chartres*. No entanto, apenas em 1915 tem seu funcionamento iniciado, sendo interrompido em 1923 e retomado em 1931. Posteriormente, dois Cursos foram instalados no Estado de São Paulo, um no Instituto Mackenzie (1929) e o outro junto à Prefeitura Municipal da cidade de São Paulo (1936), baseados na corrente norte americana da *Columbia University* (ALMEIDA; BAPTISTA, 2013).

Os cursos do Rio de Janeiro (RJ) e São Paulo (SP) não possuíam apenas raízes diferenciadas de acordo com suas correntes teóricas, mas também, nas disciplinas escolares. O curso no RJ tinha como concepção um escopo humanístico, na medida em que o curso em SP apresentava uma visão basicamente técnica (QUADRO 1).

QUADRO 1 - Disciplinas da Biblioteconomia no Rio de Janeiro x São Paulo

ANO	RIO DE JANEIRO (BN)	ANO	SÃO PAULO
1915	Bibliografia Paleografia e Diplomática Referência	1929	Catálogoção Classificação Organização de Bibliotecas
1931	Bibliografia Paleografia e Diplomática História da Literatura Iconografia e Cartografia	1941-1942	Catálogoção Classificação Bibliografia História do Livro Organização de Bibliotecas
1944	Organização e Administração de Bibliotecas Catálogoção Classificação Bibliografia e Referência História do Livro e das Bibliotecas História da Literatura (aplicada à Bibliografia) Noções de Paleografia	1943-1959	Catálogoção Classificação Bibliografia Organização de Bibliotecas História do Livro e Paleografia
1962	Técnica de Referência Bibliografia Geral Catálogoção e Classificação Organização e Administração de Bibliotecas História do Livro e das Bibliotecas Organização e Técnicas de Documentação Literatura e Bibliografia Literária Introdução à Cultura Histórica e Sociológica Reprodução de Documentos Paleografia	1960-1961	Catálogoção Classificação Referência e Bibliografia História do Livro Paleografia Organização e Administração de Bibliotecas Seleção de Livros Introdução à Cultura Artística Introdução à Cultura Filosófica Introdução às Ciências Sociais Documentação

Fonte: Castro (2000, p. 105).

É fácil observar que desde muito cedo as atividades de organização e representação do conhecimento já se configuravam em disciplinas acadêmicas tais como a bibliografia (advinda da Documentação) e a catalogação e classificação (Biblioteconomia americana). Os currículos sofreram modificações ao longo do tempo, e a padronização das disciplinas deu-se por meio do primeiro Currículo Mínimo, estabelecido em 1962, a partir da obrigatoriedade de que os diplomas de Biblioteconomia fossem registrados na Diretoria de Ensino Superior do Ministério da Educação e Cultura. Com a necessidade de atualização dos currículos, em 1982 há uma nova reestruturação e o Conselho Federal de Educação, em parceria com a Associação Brasileira de Escolas e Biblioteconomia e Documentação (ABEBD) e professores de diversos cursos de Biblioteconomia do país estabelecem o 2º Currículo Mínimo (Quadro 2) para a formação em Biblioteconomia, com a reorganização das disciplinas em três grupos: fundamentação geral, matérias instrumentais e formação profissional (ALMEIDA; BAPTISTA, 2013). Há nesse momento uma junção das correntes americanas e europeias, em benefício de uma formação abrangente (QUADRO 2).

QUADRO 2 - Currículo Mínimo 1962 X Currículo Mínimo 1982

Currículo Mínimo de 1962	Currículo Mínimo de 1982
	<i>Matérias de Formação Geral</i>
	1. Comunicação
1. Introdução aos estudos históricos e sociais	2. Aspectos sociais, políticos e econômicos do Brasil contemporâneo
2. História da arte	
3. Evolução do pensamento filosófico e científico	3. História da Cultura
4. História da literatura	
	<i>Matérias Instrumentais</i>
	4. Lógica
	5. Língua portuguesa e literatura da língua portuguesa
	6. Língua estrangeira moderna
	7. Métodos e técnicas de pesquisa
	<i>Matérias de Formação Profissional</i>
5. Documentação	8. Informação aplicada à Biblioteconomia
6. História do Livro e das Bibliotecas	9. Produção dos registros do conhecimento
	10. Formação e desenvolvimento de coleções
7. Catalogação e classificação	11. Controle bibliográfico dos registros do conhecimento
8. Bibliografia e referência	12. Disseminação da informação
9. Organização e administração de bibliotecas	13. Administração de bibliotecas
10. Paleografia	

Fonte: Almeida; Baptista, 2013, p. 7 (Adaptado de MÜLLER, 1988).

De acordo com Müeller (1985), o Instituto Brasileiro de Bibliografia e Documentação — IBBD, hoje Instituto Brasileiro de Informação em Ciência e Tecnologia - IBICT, também teve significativa influência no conteúdo dos cursos da época.

O IBBD foi criado com muito apoio da UNESCO, que para aqui trouxe dois peritos em Documentação, Herbert Coblans (1953) e Zeferino Ferreira Paulo (1954). Os esforços de Coblans, no sentido de difundir aqui métodos e técnicas de Documentação, tiveram efeito de longo prazo, com a inclusão da matéria Documentação no currículo mínimo que seria aprovado em 1962 (MÜELLER, 1985, p. 05).

Müeller (1985, p.07) ressalta que: “As disciplinas dominantes nos currículos plenos eram as relacionadas com as matérias do currículo mínimo, na proporção dada sobre o total de disciplinas oferecidas”. Nesse processo quantitativo, os números indicavam cursos fortemente estruturados em organização de acervos e da instituição biblioteca, no qual as disciplinas dessas atividades ocupavam 46.1% dos currículos plenos, assim distribuídas: Catalogação e Classificação - 22.8%; Bibliografia e Referência - 12.7% (com predominância da Bibliografia); Administração e Organização de Bibliotecas — 10.6%.

Com o crescimento do número de cursos de graduação, a formação de pessoal especializado também desponta e a produção na área científica tende a acompanhar este desenvolvimento. Danuello e Guimarães (2005) evidenciam a produção científica entre docentes nos cursos de Biblioteconomia do país e apontam o expressivo percentual de publicações em ORC em relação às demais áreas (QUADRO 3).

QUADRO 3 – Produção docente na área de Organização e Representação do Conhecimento

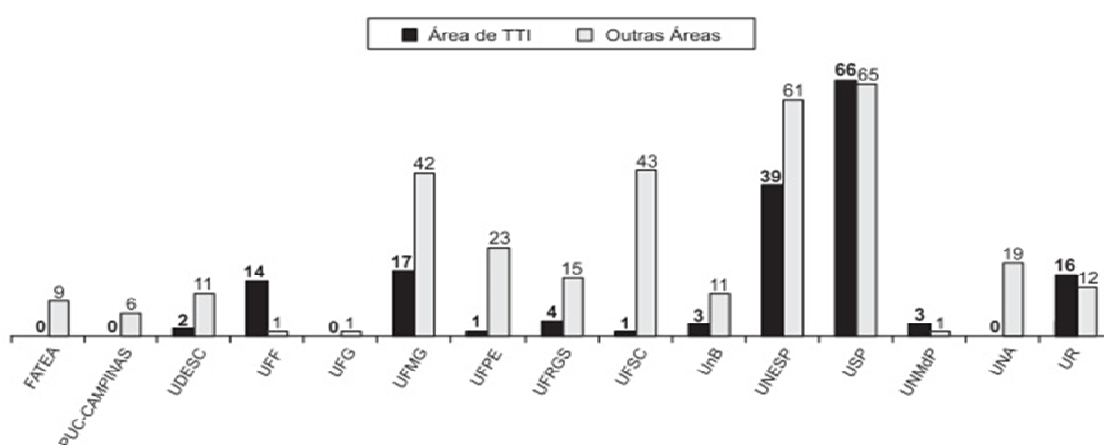


Figura 1. Distribuição da produção docente.

Nota: **FATEA**: Faculdades Integradas Teresa D'Ávila; **PUC-CAMPINAS**: Pontifícia Universidade Católica de Campinas; **UDESC**: Universidade do Estado de Santa Catarina; **UFF**: Universidade Federal Fluminense; **UFG**: Universidade Federal de Goiás; **UFMG**: Universidade Federal de Minas Gerais; **UFPE**: Universidade Federal de Pernambuco; **UFRGS**: Universidade Federal do Rio Grande do Sul; **UFSC**: Universidade Federal de Santa Catarina; **UnB**: Universidade de Brasília; **UNESP**: Universidade Estadual Paulista; **USP**: Universidade de São Paulo; **UNMdP**: Universidad Nacional de Mar del Plata; **UNA**: Universidad Nacional de Asunción; **UR**: Universidad de la Republica.

Fonte: Danuello; Guimarães, 2005.

Retomando o contexto histórico da institucionalização científica da ORC no Brasil, outro importante fator e, reconhecidamente decisivo, foi a criação dos cursos de pós-graduação. O primeiro curso de mestrado estabeleceu-se no IBBD, no ano de 1970, com o nome de Mestrado em Ciência da Informação. Os programas de pós-graduações em Ciência da Informação desenvolveram-se ao longo de 40 anos e hoje contam com cursos de Mestrado, Doutorado e Pós-Doutorado, constituindo um forte indicativo de institucionalização por possuírem em seus núcleos a essência da pesquisa e promoverem a divulgação científica. Hoje somam 14 Programas de Pós-Graduação, distribuídos por nove estados brasileiros concentrados nas regiões Nordeste, Centro Oeste, Sul e Sudeste do país. Destes, nove possuem doutorado e mestrado (UFBA, UFPB, UFPE, UFF, UFSC, USP, Unesp, UFRJ, UnB, UFMG), dois oferecem apenas mestrado (UFPE, UEL) e dois ofertam mestrados profissionais (UDESC e UNIRIO).

Dos 14 Programas, 11 apresentam linhas de pesquisa em ORC e apenas os Programas da UFBA; UDESC, e um dos mestrados da UNIRIO, não contemplam tais estudos. Esses dados demonstram o alto nível de institucionalização social da área de ORC no núcleo dos Programas de Pós-Graduação que, conseqüentemente, promoverão o avanço da institucionalização cognitiva e formação especializada nessa área de pesquisa. A presença acentuada da ORC entre as linhas de pesquisas é também um indicativo no fortalecimento de pesquisas e produção científica, uma vez que, em grande maioria, as teses e dissertações contribuem para fomentar investigações e formar profissionais especializados. No âmbito da comunicação científica, através de periódicos científicos, a ORC não dispõe de uma publicação especializada, no Brasil. Suas contribuições estão distribuídas entre os periódicos da área de CI, que são mantidos em parte pelos Programas de Pós-Graduação, Departamentos de Graduação, Institutos de pesquisa, Associações de Classes e Pós-Graduações Multidisciplinares. No Brasil, o maior número de publicações em ORC concentra-se no interior do GT2, no ENANCIB, e nas atuais publicações resultantes do Congresso ISKO-Brasil.

O Conselho Nacional de Desenvolvimento Científico e Tecnológico CNPq mantém o Diretório dos Grupos de Pesquisa, que se constitui como um inventário dos grupos de pesquisa científica e tecnológica ativos no país. Os registros, abertos ao acesso público, apresentam informações sobre a composição dos recursos humanos de cada grupo, suas linhas de pesquisa em andamento, as áreas de especialidade do conhecimento em que atuam, os setores de aplicação, produção científica e parcerias entre grupos e instituições. Neste âmbito, atualmente a CI possui 226 cursos cadastrados, dos quais, 56 Grupos apresentam Linhas de

pesquisa em ORC, respondendo por 25% dos grupos cadastrados no Diretório. O pioneirismo em investigação de temas em ORC também está presente na constituição dos Grupos de Pesquisa. Dos 226 Grupos cadastrados no Diretório do CNPq, 25 Grupos possuem mais de 15 anos de existência e destes, 4 Grupos possuem linhas em ORC atuando há aproximadamente 19 anos, com destaque para os Grupos: Grupo TEMMA/USP (1986); Organização do conhecimento/IBICT (1992); Análise Documentária/ Unesp (1993) e Formação e Atuação Profissional em Organização da informação/Unesp (1995).

Como exposto anteriormente, o ENANCIB, através do GT2 e a ISKO-Brasil são os principais eventos promotores de divulgação científica na área de ORC. Whitley (1974) aponta que a comunicação científica informal, realizada através de eventos científicos proporciona não só contatos pessoais importantes para obtenção de informação e legitimação de trabalhos e pesquisas, como também, permite necessariamente a validação confiável, realizada por colegas, através de mecanismos de filtragem que proporcionarão consenso na definição de um fenômeno, aplicação de técnicas apropriadas e resultados seguros, indicando altos níveis de institucionalização cognitiva e, conseqüentemente, o fortalecimento da institucionalização social. As 14 edições do ENANCIB, somaram ao longo dos anos, 2.432 trabalhos publicados e destes, 375 trabalhos, equivalentes a 13% do total de todas as publicações, correspondem à produção destinada ao GT2.

Outro significativo evento, que reflete números positivos para a área de ORC, é o Congresso da ISKO no Brasil. A trajetória de sedimentação e de interlocução científica da ISKO-Brasil, tem início em 2011, por ocasião do I Congresso ISKO-Brasil. O evento obteve uma excelente receptividade entre os pesquisadores, uma vez que, sua avaliação “apontou para perspectivas bastante promissoras no que se refere a pesquisas conjuntas, diálogos interdisciplinares e visibilidade científica, sinalizando a necessidade de sua continuidade e verticalização em edições futuras” (DODEBEI, GUIMARÃES, 2013, p. 13). Em 2013 foi realizado o II Congresso ISKO-Brasil, reforçando a sedimentação de reuniões com pesquisadores brasileiros e promovendo mais um espaço para interlocução e divulgação entre pesquisas e pesquisadores no âmbito nacional e internacional. O evento pautou-se na diversidade institucional, bem como na diversidade de especialidade e presença de pesquisadores de outros países.

Isto parece confirmar que, mediante a institucionalização, a área de ORC responde de forma positiva à criação e manutenção de estruturas formais que demarcam os estudos e resultados da estrutura científica no âmbito brasileiro. É importante ressaltar que o GT2 é reconhecidamente o espaço de divulgações científicas de pesquisas em ORC e possui

dimensões amplas e a ISKO “possui uma especificidade própria ao que se alia ser um foro científico de pesquisadores plenos (com titulação mínima de doutor)” (GUIMARÃES, DODEBEI, 2012, p. 17). A constituição da ISKO- Brasil já realizou duas edições e o evento é bienal e seus anais não correspondem apenas à publicação dos trabalhos aprovados, eles passaram a constituir, no Brasil, um livro digital para cada edição do evento.

Com um efetivo crescimento da comunidade científica nos últimos 20 anos, pesquisadores brasileiros têm uma atuante participação nos capítulos internacionais e regionais da ISKO. Conforme dados do último Capítulo, realizado na Índia, em 2012, entre os 51 trabalhos publicados nos anais, 15 eram de pesquisadores brasileiros, representando uma efetiva parcela de 29,4%, das publicações. Não obstante, o Capítulo Regional mais recente da ISKO Espanha e Portugal, realizado na cidade do Porto, em Portugal, no ano de 2013, reuniu 94 publicações em seus anais, dos quais, 60 trabalhos eram de pesquisadores brasileiros, correspondentes a 63,8% do montante. Para além desses dados, a esfera brasileira na *International Society for Knowledge Organization*, está contemplada com a participação do Professor José Augusto Chaves Guimarães (Universidade Estadual Paulista, Marília - UNESP) e professor Renato Rocha Souza, (Universidade federal de Minas Gerais – UFMG) entre os Conselheiros da Diretoria Executiva e Conselho Editorial.

4 CONSIDERAÇÕES FINAIS

A análise histórica da ORC permite confirmar que a institucionalização científica da área mostra-se em franca evolução, acompanhando a grande área da Ciência da Informação, no Brasil. Para Whitley (1974), a institucionalização desenvolve-se em âmbitos social e cognitivo, e estes são medidos através de critérios que indicam os níveis, alto ou baixo, de cada uma dessas vertentes. Procurou-se observar nesse levantamento, a institucionalização da ORC, em CI, no Brasil, de forma globalizada, evidenciando sua constituição e desenvolvimento que a institucionalizaram como área científica. Preliminarmente, a institucionalização social da área é apontada na constituição das pós-graduações que em maioria ofertam linhas de pesquisa em ORC e na formação dos grupos de pesquisa que conduzem investigações específicas. Os eventos, como o ENANCIB, através do GT2 e a ISKO-Brasil demarcam um campo próprio para discussões em ORC e a reunião de pesquisadores. Tais conclusões demonstram, conseqüentemente, que sua institucionalização cognitiva também apresenta bons níveis, uma vez que Whitley (1974) pondera que é pouco provável que ocorra um baixo nível de institucionalização cognitiva em uma área que

apresente um elevado nível de institucionalização social. Os expressivos números de produção científica, na área de ORC, dentro e fora do país já confirmam essa evidência.

Esta pesquisa é parte de uma investigação em andamento, que tende a desdobrar-se para obtenção individualizada e pormenorizada dos níveis de institucionalização social e cognitiva da ORC, à medida que se espera apresentar contribuições para novas investigações e para o aprofundamento de estudos dessa natureza, que possam identificar a evolução e as tendências de áreas específicas e nucleares de conhecimento para a Ciência da Informação no Brasil, como a ORC.

REFERENCIAS

- ALMEIDA, Neilia Barros Ferreira de; BAPTISTA, Sofia Galvão. Breve histórico da Biblioteconomia brasileira: formação do profissional. In: Congresso Brasileiro de Biblioteconomia, Documento e Ciência da Informação, 25., 2013, Florianópolis, SC. **Anais...** Santa Catarina: FEBAB, 2013. p. 1-12. Disponível em: <<http://portal.febab.org.br/anais/article/view/1508>>. Acesso em: 30 nov. 2013
- BOURDIEU, P. Os usos sociais da ciência: por uma sociologia clínica do campo científico. São Paulo: Ed. UNESP, 2004.
- DAHLBERG, I. Knowledge organization: a new science? **Knowledge organization**, v. 33, n. 1, 2006.
- DODEBEI, Vera Lúcia; GUIMARÃES, José Augusto Chaves. Introdução. In: DODEBEI, Vera Lúcia; GUIMARÃES, José Augusto Chaves. (Orgs.) II CONGRESSO BRASILEIRO EM REPRESENTAÇÃO E ORGANIZAÇÃO DO CONHECIMENTO, 2013. Rio de Janeiro. Anais... Rio de Janeiro, 2013, On-Line.
- EGAN, M. E.; SHERA, J. H. **Prolegomena to bibliographic control**. Journal of Cataloging and Classification, v.5, n.2, p.17-19, 1949.
- GUIMARÃES, José Augusto Chaves. A dimensão teórica do tratamento temático da informação e suas interlocuções com o universo científico da International Society for Knowledge Organization (ISKO). **Revista Ibero-americana de Ciência da Informação**, n.1, p. 77-99, 2008. Disponível em: <<http://seer.bce.unb.br/index.php/RICI/article/viewArticle/2761>>. Acesso em: 20 abril 2013.
- GUIMARÃES, José Augusto Chaves; PINHO, Fabio, Assis.; FERREIRA, Gustavo, Melett. Relações teóricas da organização do conhecimento com as abordagens de catalogação de assunto, indexação e análise documental. **Scire** (Zaragoza), v. 18, p. 33-43, 2012. Disponível em: <<http://ibersid.eu/ojs/index.php/scire/article/viewFile/3971/3704>>. Acesso em: 21 set. 2013.
- HJØRLAND, Birger. Domain analysis in Information Science: Eleven approaches – traditional well as innovative. **Journal of Documentation**, v.58, n.4, p.422-462, 2002.
- LÓPEZ YEPES, José. **Teoría de la Documentación**. Pamplona: Ediciones Universidad de Navarra (EUNSA), 1978.

MÜELLER, Suzana Pinheiro Machado. O ensino de biblioteconomia no Brasil. **Ciência da Informação**, v. 14, n. 1, p. 3-15, jan./jun. 1985. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/1452>> . Acesso em: 23 nov. 2013.

MURGUIA, Eduardo Ismael; SALES, Rodrigo de. Práticas Discursivas na Organização Do Conhecimento: Otlet, Bliss e Isko. In: RIBEIRO, Fernanda; CERVEIRA, Maria Elisa. (Org.). **Informação e/ou Conhecimento**: as duas faces de Jano. Porto: Faculdade de Letras do Porto; CETAC/ MEDIA, 2013, p. 421-436. Disponível em: <<https://ocs.letras.up.pt/index.php/ISKO/ISKO>>. Acesso em: 20 dez 2013.

ORTEGA, Cristina Dotta. Relações históricas entre Biblioteconomia, Documentação e Ciência da Informação. **DataGramZero** - Revista de Ciência da Informação, v.5, n.5, out., 2004. Disponível em: < http://www.dgz.org.br/out04/Art_03.htm >. Acesso em: 10 set. 2013.

OTLET, Paul. *Traité de documentation: le livre sur le livre: théorie et pratique*. Bruxelles: Mundaneum, 1934.

PEDROSO IZQUIERDO, Evelyn. Breve historia del desarrollo de la Ciencia de la Información. **ACIMED**, v. 12, n. 2, abr. 2004 . Disponível em: <http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1024-94352004000200007&lng=es&nrm=iso>. Acesso em: 02 maio. 2013

PINHO, Fábio Assis. **Fundamentos da Organização e Representação do Conhecimento**. Recife: Editora Universitária da UFPE, 2009.

RIBEIRO, Fernanda; CERVEIRA, Maria Elisa. Apresentação. In: RIBEIRO, Fernanda; CERVEIRA, Maria Elisa. (Org.). **Informação e/ou Conhecimento**: as duas faces de Jano. Porto: Faculdade de Letras do Porto; CETAC/ MEDIA, 2013, p. 15. Disponível em: <<https://ocs.letras.up.pt/index.php/ISKO/ISKO>>. Acesso em: 20 dez 2013.

SOUZA, Francisco das Chagas de. **Organização do conhecimento na sociedade**. Florianópolis: NUP, 1998. (Coleção Cadernos CED).

WHITLEY, Richard. Cognitive and social institutionalization of scientific specialities and research areas. In: WHITLEY, Richard. **Social processes of scientific development**. London: Routledge and Kegan, 1974. p. 69-95.

ZANDONADE, Tarciso. **As implicações da epistemologia social para uma teoria da recuperação da informação**. Tese (Doutorado em Ciência da Informação e Documentação). Universidade de Brasília - UNB, Brasília, 2003.

PROPOSTA DE MODELAGEM CONCEITUAL A PARTIR DA IMPLEMENTAÇÃO DA TEORIA DA CLASSIFICAÇÃO FACETADA DE RANGANATHAN EM BANCO DE DADOS

PROPOSAL FOR A MODELLING CONCEPT FROM THE IMPLEMENTATION OF THE RANGANATHAN THEORY OF CLASSIFICATION FACETED IN DATABASE

Márcio Bezerra Da Silva
Zeny Duarte de Miranda

Resumo: Apresenta uma proposta de implementação da Teoria da Classificação Facetada (TCF) de Ranganathan na modelagem conceitual em Banco de Dados (BD). Discute a TCF a partir das categorias fundamentais Personalidade, Matéria, Energia, Espaço e Tempo (PMEST). Debate a formação de assuntos e isolados: *Dissecação*, *Laminação*, *Desnudação*, *Reunião/Agregação* e *Superposição*. Objetiva-se, de forma maior, a partir de um percurso metodológico formado pela Pesquisa Aplicada, Exploratória e Bibliográfica, investigar a implementação da TCF de Ranganathan na modelagem conceitual de dados, visando à organização da informação; e de maneira específica, apresentar um modelo de organização da informação pelos métodos *Dissecação* e *Desnudação*, e proporcionar uma proposta de organização da informação pela técnica *Laminação*. Coleta, como resultados da pesquisa, que o modelo de mapa categorial proveniente da *Laminação* apresentou uma redução na quantidade de conceitos, as opções de facetação também diminuiriam ao ponto da exclusão dos focos e conceitos que possuem algum tipo de relação entre si poderão não ser oferecidos ao usuário no momento da busca. Resulta também que o modelo gerado pela *Desnudação*, ao oferecer diferentes visões (focos) sobre a mesma ocorrência apresentada nos resultados de busca, teoricamente, permite ao usuário maior autonomia de navegação. Conclui-se que a *Laminação* apresentou uma redução na quantidade de conceitos, inferindo resultados mais precisos, ao passo que a *Desnudação*, teoricamente, permite maior autonomia de navegação ao usuário, uma vez que oferece diferentes visões sobre a mesma ocorrência apresentada nos resultados de busca, propriedade cada vez mais adotada na *Web*.

Palavras-chave: Teoria da Classificação Facetada de Ranganathan. Modelagem Conceitual. Categorias Fundamentais. Formação de assuntos e isolados.

Abstract: Presents a proposal for the implementation of the Ranganathan Theory of Faceted Classification (TCF) in conceptual modeling in the database (DB). Discusses the TCF from the fundamental categories Personality, Matter, Energy, Space and Time (PMEST). Debate the formation of subjects and isolated: Dissection, Lamination, Denudation, Loose Assemblage and Superimposition. Objective is, in a greater form, according to a methodological path formed by the Research Technique Applied, Exploration and Bibliographic, investigate the implementation of TCF Ranganathan in conceptual data modeling, aimed at organizing information; and specific objective present a model of organizing information by dissection and stripping methods, and provide a proposal for organizing information by lamination technique. How have the results of the research the model from the categorical map Laminating decreased the amount of concepts, the facing options also decreased to the point of excluding foci and concepts that have some kind of relation may not be offered the user at the time of search. It is also apparent that the model generated by Denudation, by offering different visions (focus) on the same event displayed in search results, theoretically, allows the user greater autonomy navigation. It is concluded that the lamination showed a reduction in the amount of concepts, inferring more accurate results, while denudation theoretically allows greater autonomy navigation to the user, since it offers

different views of the same event displayed in search results owned increasingly adopted the Web

Keywords: Faceted Classification Theory of Ranganathan. Conceptual Modeling. Fundamental categories. Formation of subjects and isolated.

1 INTRODUÇÃO

Na atual sociedade a informação assume valor essencial ao cotidiano dos indivíduos. Quanto mais amplia-se o seu valor, maior é a necessidade de organizá-la, especialmente pelos recentes adventos tecnológicos, entre os quais destacam-se os computadores e a Internet. O contexto apresentado é reflexo da necessidade humana em organizar o seu conhecimento, exigindo que os modelos para este fim sejam aperfeiçoados e/ou criados.

A necessidade de organizar o conhecimento humano transcendeu o espaço físico, chegando aos ambientes digitais. A *Web* mostra-se a cada dia mais dinâmica, intuitiva, ao passo que oferece recursos para a atual crescente produção de informação. Neste sentido, o êxito na recuperação de informação padece de frustrações, gerando a ideia de que o problema está na ferramenta de busca, entretanto, estudiosos como Campos (2001), Lima (2004), Miranda (2005), o Da Silva (2011) e Pontes (2013) permitem o vislumbre de que a real dificuldade encontra-se na etapa anterior do sistema, ou seja, na fase de organização.

A discussão aqui iniciada possui raízes na Ciência da Informação (CI), área do conhecimento que segundo Capurro (2003) busca a produção, seleção, organização, interpretação, armazenamento, recuperação, disseminação, transformação e uso da informação. Saracevic (1996) complementa afirmando que a CI é uma ciência formada pela Biblioteconomia, Matemática, Lógica, Linguística, Psicologia, Ciências Cognitivas, Ciência da Computação (CC), Comunicação entre outras.

Conforme as definições dos autores supracitados, a presente pesquisa, iniciada em 2009, encontra-se na convergência entre a CI e a CC a partir do desenvolvimento de um Banco de Dados (BD), e que fez uso da Teoria da Classificação Facetada (TCF) para organizar e recuperar a informação de certa Instituição promotora de cursos profissionalizantes, mais especificamente do seu Setor Pedagógico.

A TCF foi criada pelo indiano Shialy Rammarita Ranganathan (1897-1972), sistema que permite a estruturação do conhecimento através da organização de seus conceitos e da criação de relacionamentos, permitindo o mapeamento de uma área de assunto, bem como ocorrem em BD. Além disso, a classificação facetada possibilita a inclusão de novos conceitos sem que isto altere a estrutura do sistema, determinando assim a sua escolha de implementação.

Apesar da TCF surgir com fins de organização de acervos físicos na Índia, seus pressupostos teóricos se enquadram na premissa apresentada, ou seja, de que os sistemas de organização do conhecimento (SOC) sejam aperfeiçoados e/ou criados, neste caso, adaptado para ambientes digitais.

Intitulado de *Sistema Facetado* e apresentada suas etapas de construção¹⁷⁸ nas edições XI, XII e XIII do Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB), conforme as diferentes fases da modelagem de dados do BD, a organização das informações no sistema ficou por conta das técnicas, de formação de assuntos e isolados, chamadas *Dissecação (Dissection)*, o qual permitiu identificar o assunto básico do sistema, e *Desnudação (Desnudation)*, que permitiu representar o núcleo específico do assunto básico e das ideias isoladas.

No total, a TCF apresenta a “Dissecação, Laminação, Desnudação, Reunião/Agregação e Superposição” (RANGANATHAN, 1967, p. 351) como modelos de representação temática, considerados preliminares para a formação de assuntos e isolados. Neste sentido, vislumbrando a utilização da técnica de *Laminação* como teste para a atual escrita da tese de doutoramento do presente autor, por meio dos documentos/materiais utilizados nos cursos de Informática da Instituição solicitadora do BD, este artigo objetiva, de forma ampla, investigar a implementação da TCF de Ranganathan na modelagem conceitual de dados, visando à organização da informação. De maneira específica almeja-se apresentar um modelo de organização da informação pelos métodos *Dissecação* e *Desnudação*, e proporcionar uma proposta de organização da informação pela técnica *Laminação*.

2 A TEORIA DA CLASSIFICAÇÃO FACETADA DE RANGANATHAN

Inúmeros são os sistemas de classificação desenvolvidos ao longo da história humana. Baseados nos modelos de *Classificação Filosófica do Conhecimento* foram criados os intitulos sistemas de classificação bibliográficos. Como enfoque deste artigo, destaca-se o código cunhado em 1933 pelo matemático e bibliotecário Ranganathan (1892-1972), chamado *Colon Classification*, também conhecido como Classificação de Dois Pontos. Considerado o pai da biblioteconomia na Índia e o pai da moderna teoria da classificação, Ranganathan apresentou uma teoria com o objetivo de garantir uma sequência considerada útil dos livros nas estantes, já que a preocupação principal era com a localização física dos livros na biblioteca em relação ao tema central abordado na obra.

¹⁷⁸ Artigos produzidos ao longo do desenvolvimento da dissertação de mestrado intitulada “A Teoria da Classificação Facetada na Modelagem de Dados em Banco de Dados Computacionais” (2011).

Em sua teoria, Ranganathan faz uso de categorias fundamentais, elementos que refletem manifestações de acordo com a área do conhecimento classificada. Tais manifestações, também chamadas de facetas, possibilitam agrupar os assuntos com a classe principal e com os que tenham o mesmo tipo de relacionamento entre si. Assim, as facetas formam uma espécie de lista de termos relacionados, contendo características dos seus respectivos assuntos. Faceta é “um termo genérico usado para denotar algum componente – pode ser um assunto básico ou um isolado – de um assunto composto, tendo, ainda, a função de formar renques, termos e números” (RANGANATHAN, 1967, p. 88). Barbosa (1972, p. 75, grifo do autor) compreende faceta “como uma lista de termos mantendo entre si as mesmas amplas relações com a classe que lhes deu origem, ou então, ‘como um conjunto de termos produzidos pela aplicação de um amplo *princípio de divisão*”. Por sua vez, Miranda (2005, p. 132) resume a discussão afirmando que faceta é “um termo genérico usado para designar qualquer componente de um assunto composto”. Em suma, as categorias fundamentais são formadas pela Personalidade, Matéria, Energia, Espaço e Tempo (PMEST).

2.1 Categorias Fundamentais (PMEST)

A primeira categoria fundamental chama-se Personalidade [P], considerada a essência de um determinado assunto, como também os objetos de estudo de uma determinada disciplina, tipos, entidades etc. Conforme Vickery (1980, p. 212), é possível citar como exemplo da categoria Personalidade [P]: “bibliotecas, números, equações, comprimentos de ondas de irradiação, obras de engenharia, substâncias químicas, organismos e órgãos, adubos, religiões, estilos de arte, línguas, grupos sociais, comunidades”.

Matéria [M] representa manifestações, complementos, substâncias, que constituem as coisas, os objetos. Esta categoria é classificada em material e propriedade. Por exemplo, na biblioteconomia tem-se livros (material) raros (propriedade).

Energia [E] ou *ação* é a manifestação de um verbo, de uma ação, como processos, técnicas, atividades etc. Por exemplo, na biblioteconomia, serviços como catalogação, indexação e classificação são exemplos da categoria Energia (ação).

Espaço [S] é uma divisão geográfica, uma manifestação de lugar, onde ocorre determinado evento, como cidades, superfícies em geral entre outros. Exemplificando, tem-se a cidade de Brasília (DF) como local onde existe um dos cursos de biblioteconomia do país.

Por fim, a categoria fundamental Tempo [T] é uma a divisão cronológica, uma manifestação de ideias em determinado tempo comum, como séculos, anos, meses, dias etc.

Como exemplo, cita-se o ano de 1965, momento que marcou a fundação do curso de biblioteconomia na Universidade de Brasília (UnB).

Das categorias da mnemônica PMEST surgem as facetas. Segundo Ranganathan (1967), categorias são as classes mais gerais, que podem ser formadas e empregadas para reunir outros conceitos. Cada faceta é formada por um número mínimo de subdivisões, denominadas *focos*. Após a faceta principal, surgem os *termos* que são as representações verbais dos conceitos em uma linguagem natural.

2.2 Formação de Assuntos e Isolados

Além de determinar os elementos que caracterizam a TCF, em especial suas facetas, Ranganathan (1967, p. 351) também estipulou cinco caminhos para a formação dos assuntos e isolados, e conseqüentemente ordenar as facetas, focos e termos: “Dissecação, Laminação, Desnudação, Reunião/Agregação e Superposição”.

A *Dissecação (Dissection)* propõe dividir o universo em partes coordenadas em um mesmo nível, quantas vezes for necessário, criando para cada parte o seu próprio universo. Cada uma dessas partes é chamada de *Lâmina* e esta pode representar um assunto básico ou um isolado, conforme pode ser observado no QUADRO 1.

QUADRO 9 – Exemplo de Dissecação.

Universo de Assunto Básico	Universo de Isolado: Plantas Agrícolas
Botânica	Plantas Forrageiras
Agricultura	Plantas Alimentícias
Zoologia	Plantas Estimulantes

Fonte: Adaptado de Campos (2001, p. 66).

O modo da *Laminação (Lamination)* propõe-se a superposição de uma faceta com outra, constituindo-se camadas de assuntos básicos e ideias isoladas. Além disso, as ideias isoladas, quando combinados, formam assuntos compostos, segundo exemplo apresentado no QUADRO 2.

QUADRO 2 – Exemplo de Laminação.

Assunto Básico	Ideia(s) Isolada(s)	Assunto(s) Composto(s)
Agricultura	Milho	Agricultura do Milho
Agricultura	Java	Agricultura do Milho em Java

Fonte: Adaptado de Campos (2001, p. 67).

Já a *Desnudação* (*Denudation*), também chamada de *Desfolhamento*, provoca uma diminuição progressiva da extensão e um aumento da profundidade de um assunto básico ou de uma ideia isolada, permitindo tanto a formação de cadeias, através de resultados sucessivos deste caminho/processo, quanto representar o núcleo específico de um assunto básico ou de uma ideia isolada. Observe um exemplo de *Desnudação* no QUADRO 3.

QUADRO 3 – Exemplo de *Desnudação*.

Assunto Básico
Filosofia
Lógica
Lógica Dedutiva

Fonte: Adaptado de Campos (2001, p. 67).

Quanto a *Reunião/Agregação* (*Loose Assemblage*), refere-se à combinação do assunto básico ou composto com ideias isoladas, formando um assunto complexo ou uma ideia isolada complexa, conforme pode ser observado no QUADRO 4.

QUADRO 4 – Exemplo de *Reunião/Agregação*.

Assunto Básico 1	Assunto Básico 2	Assunto Complexo
Ciência Política	Economia	Relação Geral entre Ciência Política e Economia
Vertebrados	Invertebrados	Diferença entre Vertebrados e Invertebrados

Fonte: Adaptado de Campos (2001, p. 67).

Por fim, a *Superposição* (*Superimposition*), também chamada de *Sobreposição*, permite a conexão de duas ou mais ideias isoladas que pertencem ao mesmo universo de ideias isoladas. A ideia isolada resultante desta superposição é chamada de *ideia isolada superposta* ou *ideia isolada composta*. Este caminho é diferente do adotado pela *Laminação*, que permite a ligação entre isolados de universos diferentes. Observe um exemplo de *Superposição* no QUADRO 5.

QUADRO 5 – Exemplo de *Superposição*.

Ideia Isolada	Característica Assunto	Característica Habilidade Retórica	Ideia Isolada Superposta
Professor	Química	Brilhante	Professor de Química Brilhante
Professor	Química	Medíocre	Professor de Química Medíocre
Professor	Zoologia	Brilhante	Professor de Zoologia Brilhante
Professor	Zoologia	Medíocre	Professor de Zoologia Medíocre

Fonte: Adaptado de Campos (2001, p. 67).

Todos os conceitos usados para a formação de assuntos e isolados devem estar organizados no sistema de classificação. Estes se organizem de maneira hierárquica (gênero-espécie/todo-parte), em *Array* (renques) e *Chain* (cadeias).

Conforme Ranganathan (1967) renques são conjuntos formados a partir de uma única característica de divisão, formando séries horizontais, e cadeias são séries verticais de conceitos em que cada conceito tem uma característica a mais ou a menos, conforme a cadeia seja ascendente ou descendente. No quadro 6, é possível observar um modelo de organização por renque e outro por cadeia, tomando como exemplo “Flores Ornamentais”. No caso do renque, Rosa e Margarida são tipos de Flores Ornamentais. Já na organização por cadeia, Rosa é um tipo de “Flor Ornamental” que, por sua vez, é um tipo de Flor.

QUADRO 6 – Exemplo de Renque e Cadeia.

Renque	Cadeia
Flores Ornamentais	Rosa
Rosa	Flor Ornamental
Margarida	Flor

Fonte: Adaptado de Anjos (2008).

Entender esses caminhos “é fundamental, pois desta compreensão resultará a sua atuação nos processos de organização e recuperação de informação” (CAMPOS, 2001, p.66). Definida essa etapa os assuntos e isolados poderão ser relacionados de maneira ilimitada e multidimensional, levando-se em consideração as diferentes visões do mesmo objeto, ou seja, dos diferentes focos atribuídos, conforme promulga a TCF.

3 CARACTERÍSTICAS METODOLÓGICAS DA PESQUISA

Esta pesquisa possui como base discursiva a aplicação de estudos em representação e organização do conhecimento, enquanto teorias da CI, em espaços digitais, neste caso, o ambiente BD da CC. Neste sentido, do ponto de vista da natureza, trata-se de uma Pesquisa Aplicada, que “objetiva gerar conhecimentos para aplicação prática dirigidos à solução de problemas específicos. Envolve verdades e interesses locais (SILVA, MENEZES, 2001, p. 20). Quanto ao ponto de vista de seus objetivos, enquadra-se como Pesquisa Exploratória, porque é “[...] dada à descoberta de práticas ou diretrizes que precisam modificar-se e na elaboração de alternativas que possam ser substituídas” (OLIVEIRA, 2001, p. 134). Especialmente pela quantidade mínima de modelos empíricos que apresentam a TCF implementadas em ambientes digitais, com ênfase nos estudos em BD, a Pesquisa Exploratória “[objetiva] proporcionar maior familiaridade com o problema, com vistas a

torná-lo mais explícito ou constituir hipóteses” (GIL, 2002, p. 41). No que se refere ao ponto de vista dos procedimentos técnicos, adotou-se a pesquisa bibliográfica, que segundo Gil (1989, p. 71) “é desenvolvida a partir de material já elaborado, constituído principalmente de livros e artigos científicos”. No caso deste artigo, fazendo uso de livros, artigos científicos, dissertação e teses, a pesquisa bibliográfica permitiu “conhecer as diferentes formas de contribuição científica que se realizaram sobre determinado assunto ou fenômeno” (OLIVEIRA, 2001, p.119).

Para formalizar os modelos de representação e organização da informação no sistema foi delineado o uso das técnicas de formação de assuntos e isolados *Dissecação* e *Desnudação*, já testadas, e neste artigo a adoção da *Laminação*.

Como a Instituição atua no campo da Capacitação Profissionalizante, os termos extraídos dos documentos/materiais representam a realidade do Setor Pedagógico, seção responsável pela coordenação dos cursos, além da aquisição e produção dos materiais didáticos. No caso deste trabalho o enfoque manteve-se aos documentos/materiais dos cursos de Informática, objetos que subsidiaram a extração dos termos que constituíram a base funcional do BD. Deve-se salientar que diante da problemática na organização desses documentos/materiais, muitos foram adquiridos mais de uma vez, como livros e programas de computador.

4 PROPOSTAS DE ORGANIZAÇÃO DA INFORMAÇÃO PELA TCF

O desenvolvimento do *Sistema Facetado*, sendo um Sistema de Banco de Dados (SBD), ocorreu no cumprimento de três etapas chamadas de modelagens de dados: conceitual, lógica e física. No caso deste trabalho, como o enfoque enquadra-se na TCF e não na prototipagem do sistema, os resultados da pesquisa referem-se a modelagem conceitual, definida por Castro (2012, p. 8) um esquema “[...] que descreve a estrutura do banco de dados sob o ponto de [vista] do consumidor da informação (usuário) sem preocupação com detalhes físicos de armazenamento”. Ainda segundo Castro (2012, p. 10), a etapa de modelagem em debate possui “a finalidade de descrever os dados necessários para a geração das informações requeridas em um contexto analisado, numa estreita aproximação com a forma como o usuário/desenvolvedor, vê os dados e analisa sua utilidade para a geração de informação”.

Na modelagem conceitual não são levadas em consideração atribuições computacionais relacionadas à escolha, neste caso, do modelo de desenvolvimento lógico, do Sistema Gerenciador de Banco de Dados (SGBD) e da resultante de prototipagem, ou seja, o próprio SBD, aqui rotulado de *Sistema Facetado*.

A etapa de modelagem em discussão foi adaptada da estrutura conceitual do Mapa Hipertextual (MHTX) criado por Lima (2004), em sua tese, a partir da inclusão das seguintes ações: conversar com os possíveis usuários do sistema; compreender as atividades realizadas pelos usuários na Instituição; identificação dos materiais pertencentes ao Setor Pedagógico; leitura técnica dos materiais para a extração dos termos representativos; análise facetada; organização de todos os termos selecionados e suas respectivas relações.

A resultante foi um conjunto de termos indexados pela linguagem natural, ou seja, conforme são apresentados nos documentos/materiais, não passando pela tradução de qualquer vocabulário controlado. O QUADRO 7 representa a etapa analítica da TCF, que agrupa os termos sem qualquer ordenação (classificação).

QUADRO 7 – Lista de termos extraídos dos documentos/materiais de Informática.

Termos Indexados	
Administrativo	Introdução a Microinformática
Apostila	Laboratório 2
Apresentação	Laboratório de informática
Artes e cartões	Lista por curso
Auxiliar Administrativo	Livro
Avançado	Manhã
Básico	Modular
Cadastro de aluno	Nivelamento
Cadastro de professor	Noite
Capacitação	Matéria
CD-ROM	Primeiro
Cursos	Professor
Documentos	Profissionais
Educação	Propriedade
Ensino	Relatórios
Informática	Segundo
Informática Básica	Semestre
Informativo	Setor Pedagógico
Interdisciplinar	Tarde
Intermediário	Técnico
Internet	Turno

Fonte: Atualizado de Da Silva (2011, p. 110). Resultado da pesquisa, 2014.

Definidos os termos que farão parte da estrutura de funcionamento do sistema, fez-se necessário distingui-los entre assuntos e isolados, e posteriormente organizá-los em facetas. Iniciando a etapa sintética da TCF, no caso deste trabalho, a referida distinção ocorreu pela *Dissecação*, *Desnudação* e *Laminação* respectivamente, conforme informado na metodologia.

A *Dissecação* permitiu dividir o universo dos materiais e identificar que “Informática” é o assunto básico. Entretanto, buscando fidelização a realidade do Setor Pedagógico, o

assunto principal do sistema foi denominado de “Educação Profissionalizante em Informática”, também relacionado pela Instituição aos isolados “Capacitação Profissional” ou “Capacitação Profissionalizante”.

QUADRO 8 – Exemplo de resultado da *Dissecação*.

Universo de Assunto Básico	Universo de Isolado: Educação
Informática	Educação Profissionalizante
Informática	Capacitação Profissional
Informática	Capacitação Profissionalizante

Fonte: Atualizado de Da Silva (2011). Resultado da pesquisa (2014).

No caso da *Desnudação*, também foi possível representar o núcleo específico de um assunto básico, ou de uma ideia isolada dos materiais consultados. A *Desnudação* foi considerada o método mais importante para o sistema, pois a Instituição necessitava que a recuperação de informações fosse precisa, exigindo maior especificidade, de acordo com a exemplificação do QUADRO 9:

QUADRO 9 – Exemplo de resultado da *Desnudação*.

Assunto Básico
Informática
Profissionais
Professor
Professor de Informática

Fonte: Atualizado de Da Silva (2011). Resultado da pesquisa, 2014.

Por sua vez, a técnica de *Laminação* permitiu a superposição de um termo com outro, constituindo-se camadas entre o assunto básico e ideias isoladas, ambas combinadas. Além disso foi possível conjugar ideias isoladas entre si. A resultante de todas as combinações feitas é chamada de assuntos compostos, segundo exemplo apresentado no QUADRO 10.

QUADRO 10 – Exemplo de resultado da *Laminação*.

Assunto Básico	Ideia(s) Isolada(s)	Assunto(s) Composto(s)
Informática	Professor	Professor de Informática
Informática	Básica	Informática Básica

Fonte: Atualizado de Da Silva (2011). Resultado da pesquisa, 2014.

Concluída a definição do assunto geral do sistema (rotulação usada pela Instituição), assunto básico, assuntos compostos e isolados, os termos precisavam ser organizados conforme suas características, ou seja, segundo a mnemônica PMEST. Entretanto alguns conceitos receberam outras denominações, situação peculiar a modelagem conceitual que deve ser fidedigna à realidade do Setor Pedagógico. Desta forma, Personalidade [P] foi traduzida por Tipologia; Matéria [M] por Propriedade (qualidade) e Matéria (substância); Energia [E] por Ação; Espaço [S] por Lugar; e Tempo [T] obviamente manteve a rotulação original.

A ocorrência da modelagem conceitual foi denominada de *Mapa Categorical do Sistema Facetado*, estrutura básica de funcionamento do sistema, entre os módulos oferecidos aos usuários, ou seja, cadastro (catalogação), classificação e pesquisa (busca x recuperação). O primeiro mapa elaborado foi consequência da *Desnudação*.

O QUADRO 11 apresenta a categoria fundamental Tipologia [P]. Os termos “Profissionais”, “Cursos” e “Documentos” representam os focos.

QUADRO 11 – Categoria Fundamental [P].

Tipologia
Profissionais
Professor
De Informática
De Educação
Técnico
Auxiliar Administrativo
Cursos
Informática Básica
Introdução a Microinformática
Internet
Artes e Cartões
Documentos
Ensino
Lista por curso
Relatórios
Administrativo
Cadastro de Professor
Cadastro de Aluno

Fonte: Atualizado de Da Silva (2011). Resultado da pesquisa, 2014.

O QUADRO 12 apresenta a categoria fundamental Propriedade/Matéria [M]. Os termos “Propriedade”, no sentido de qualidade, e “Matéria”, enquanto substância física representa os focos.

QUADRO 12 – Categoria Fundamental [M].

Propriedade e Matéria
Qualidade
Técnico
Interdisciplinar
Informativo
Nivelamento
Básico
Intermediário
Avançado
Modular
Substância
Apostila
Livro
CD-ROM

Fonte: Atualizado de Da Silva (2011). Resultado da pesquisa, 2014.

O QUADRO 13 apresenta a categoria fundamental Ação [E]. Os termos “Propriedade”, no sentido de qualidade, e “Matéria”, enquanto substância física, representam os focos.

QUADRO 13 – Categoria Fundamental [E].

Ação
Capacitação
Cursos (por suas ações realizadas/competências)
Apresentação
Palestras

Fonte: Atualizado de Da Silva (2011). Resultado da pesquisa, 2014.

O QUADRO 14 apresenta a categoria fundamental Lugar [E]. Os termos “Informática” e “Administrativo” representam os focos.

QUADRO 14 – Categoria Fundamental [S].

Lugar
Informática
Laboratório de Informática
Laboratório 2
Administrativo
Setor Pedagógico

Fonte: Atualizado de Da Silva (2011). Resultado da pesquisa, 2014.

O QUADRO 15 apresenta a categoria fundamental Tempo [T]. Os termos “Turno” e “Semestre” representam os focos.

QUADRO 15 – Categoria Fundamental [T].

Tempo
Turno
Manhã
Tarde
Noite
Semestre
Primeiro
Segundo

Fonte: Atualizado de Da Silva (2011). Resultado da pesquisa, 2014.

Finalizada a apresentação dos resultados oriundos do mapa elaborado pela *Desnudação*, inicia-se a proposta organizacional pela *Laminação*. De forma preliminar é possível perceber a combinação entre os assuntos e isolados, extinguindo os focos.

O QUADRO 16 apresenta a categoria fundamental Tipologia [P], formada pela combinação entre o assunto básico “Informática” e os isolados “Cursos”.

QUADRO 16 – Categoria Fundamental [P].

Tipologia
Professor de Informática
Curso de Informática Básica
Curso de Informática em Internet
Curso de Informática em Artes Gráficas
Curso de Introdução a Microinformática

Fonte: Resultado da pesquisa, 2014.

O QUADRO 17 apresenta a categoria fundamental Propriedade/Matéria [M], formada pela combinação entre o assunto básico “Informática” e os isolados “Técnico”, “Básico”, “Avançado”, “Módulo (modular)”, “Apostila”, “Livro”, “CD-ROM”, “Listas dos (por) cursos”, “Relatórios”, “Cadastro de Professor” e “Cadastro de Aluno”.

QUADRO 17 – Categoria Fundamental [M].

Propriedade e Matéria
Material técnico em Informática
Informática Básica
Informática Avançada
Módulo Básico de Informática
Módulo Intermediário de Informática

Módulo Avançado de Informática
 Apostila de Informática
 Livro de Informática
 CD-ROM de Informática
 Listas dos cursos de Informática
 Relatórios dos cursos de Informática
 Cadastro de Professor de Informática
 Cadastro de Aluno de Informática

Fonte: Resultado da pesquisa, 2014.

O QUADRO 18 apresenta a categoria fundamental Ação [E], formada pela combinação entre o assunto básico “Informática” e os isolados “Capacitação”, no sentido de competências de atuação adquiridas durante o curso, e “Apresentação”, enquanto palestras, conferências, seminários etc.

QUADRO 18 – Categoria Fundamental [E].

Ação
Capacitação (curso) em Informática
Apresentação (palestra) em Informática

Fonte: Resultado da pesquisa, 2014.

O QUADRO 19 apresenta a categoria fundamental Lugar [S], formada pela combinação entre o assunto básico “Informática” e o isolado “Laboratório”.

QUADRO 19 – Categoria Fundamental [S].

Lugar
Laboratório de Informática

Fonte: Resultado da pesquisa, 2014.

O QUADRO 20 apresenta a categoria fundamental Tempo [T], formada pela combinação entre o assunto básico “Informática” e os isolados “Curso”, “Manhã”, “Tarde”, “Noite”, “Primeiro”, “Segundo” e “Semestre”.

QUADRO 20 – Categoria Fundamental [T].

Tempo
Curso de Informática de Manhã
Curso de Informática de Tarde
Curso de Informática de Noite
Curso de Informática do Primeiro Semestre
Curso de Informática do Segundo Semestre

Fonte: Resultado da pesquisa, 2014.

Observando os dois modelos de mapas categoriais, o proveniente da *Laminação* apresentou uma redução na quantidade de conceitos, entretanto as opções de facetação também diminuíram ao ponto da exclusão dos focos. Por outro lado, o modelo gerado pela *Desnudação*, ao oferecer diferentes visões (focos) sobre a mesma ocorrência apresentada nos resultados de busca, teoricamente, permite ao usuário maior autonomia de navegação, característica que vem crescendo na *Web* e mais próxima aos usuários atuais. Por fim, no caso da *Laminação*, conceitos que possuem algum tipo de relação entre si poderão não ser oferecidos ao usuário no momento da busca.

5 CONSIDERAÇÕES FINAIS

Conforme o estudo apresentado é possível perceber que a organização da informação oferece subsídios positivos aos ambientes digitais, especialmente quando direcionamentos a afirmativa aos atuais problemas de recuperação. Entende-se que o problema da recuperação da informação não relevante pelo usuário é reflexo da problemática ou inexistente organização. Neste sentido, encontram-se nas práticas da CI sistemas originalmente criados e usados em coleções de livros, em ambientes físicos. Entretanto, características desses SOC como a organicidade de assuntos e as relações entre eles são plausíveis de aplicação em ambientes digitais como, por exemplo, em BD.

O SOC que se enquadrou no panorama supracitado, oriundo das classificações bibliográficas, foi a *Colon Classification*, mais especificamente sua TCF, enfoque do presente trabalho. Os estudos sobre a TCF apresentaram semelhanças com os preceitos dos BD, pois assim como na classificação facetada, os sistemas computacionais por modelagem de dados também objetivam a estruturação do conhecimento, através da organização de seus conceitos e da criação de relacionamento, entre eles, permitindo o mapeamento de uma área de assunto e a inclusão de novos conceitos, sem que isto altere a estrutura do sistema.

A TCF contribuiu efetivamente em dois momentos da modelagem conceitual do BD, intitulado *Sistema Facetado*: na formação dos assuntos e isolados, e na definição das categorias fundamentais para o funcionamento do sistema. No caso da primeira contribuição, foram adotados os métodos de *Dissecação*, o qual permitiu identificar o assunto básico do sistema, ou seja, “*Informática*”; de *Desnudação*, que permitiu representar o núcleo específico do assunto básico e das ideias isoladas, atendendo aos anseios da Setor Pedagógico em especificar ao máximo os conteúdos dos documentos/materiais; e a técnica de *Laminação*, que permitiu a superposição de um termo com outro, constituindo-se camadas entre o assunto básico e ideias isoladas, ambas combinadas. A segunda contribuição refere-se a organização

dos termos provenientes dos métodos supracitados por meio da mnemônica PMEST, que por sua vez tiveram seus rótulos adequados à realidade do Setor Pedagógico: Personalidade [P] foi traduzida por Tipologia; Matéria [M] por Propriedade (qualidade) e Matéria (substância); Energia [E] por Ação; Espaço [S] por Lugar; e Tempo [T] manteve a rotulação original.

O uso de facetas pode ser um ponto positivo aos ambientes digitais, como os BD, pois analisar um assunto por diferentes olhares, diversificados focos, permite que a característica de certo assunto seja compreendida conforme o conhecimento, a experiência, a vivência de cada pessoa, propiciando, naturalmente, um ambiente flexível, reagrupado, multidimensional e ilimitado.

Conclui-se que a *Laminação* apresentou uma redução na quantidade de conceitos, inferindo resultados mais precisos, ao passo que a *Desnudação*, teoricamente, permite maior autonomia de navegação ao usuário, uma vez que oferece diferentes visões sobre a mesma ocorrência apresentada nos resultados de busca, propriedade cada vez mais adotada na *Web*.

REFERÊNCIAS

BARBOSA, A. P. Classificações facetadas. **Ci. Inf.** Rio de Janeiro, v.1, n. 2, p. 73-81, 1972. Disponível em: <revista.ibict.br/index.php/ciinf/article/download/1665/1271>. Acesso em 19 jul. 2014.

CAMPOS, M. L. A. **A organização de unidades do conhecimento em hiperdocumentos: o modelo conceitual como um espaço comunicacional para realização da autoria.** 2001. 190f. Tese (Doutorado em Ciência da Informação) - CNPq/IBICT-URFJ/ECO, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2001.

CAPURRO, R. Epistemologia e Ciência da Informação. In: **ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO**, 5, 2003, Belo Horizonte. Disponível em: <http://www.capurro.de/enancib_p.htm>. Acesso em: 20 de jul. 2014.

CASTRO, E. B. **Modelagem lógica de dados: construção básica e simplificada.** Rio de Janeiro: Editora Ciência Moderna, 2012.

DA SILVA, M. B. **A teoria da classificação facetada na modelagem de dados em bancos de dados computacionais.** 2011. 168f. Dissertação (Mestrado em Ciência da Informação). Curso de Pós Graduação em Ciência da Informação – PPGCI, Universidade Federal da Paraíba, João Pessoa, 2011.

GIL, A. C. **Como elaborar projetos de pesquisa.** São Paulo: Atlas, 2002.

GIL, A. C. **Métodos e Técnicas de Pesquisa Social.** 2. ed. São Paulo: Editora Atlas, 1989.

LIMA, G. A. B. **Mapa Hipertextual (MHTX): um modelo para organização hipertextual de documentos.** 2004, 199f. Tese (Doutorado em Ciência da Informação). Curso de Pós Graduação em Ciência da Informação – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Minas Gerais, 2004.

MIRANDA, M. L. C. de. **Organização e Representação do Conhecimento**: fundamentos teórico-metodológico na busca e recuperação da informação em ambientes virtuais. 2005, 353f. Tese (Doutorado em Ciência da Informação). Curso de Pós Graduação em Ciência da Informação – Convênio CNPQ/IBICT-UFRJ/ECO, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.

OLIVEIRA, S. L. de. **Tratado de metodologia científica**: Projetos de Pesquisas, TGI, TCC, Monografias, Dissertações e Teses. São Paulo: Pioneira, 2001.

PONTES, F. V. **Organização do conhecimento em bibliotecas digitais de teses e dissertações**: uma abordagem baseada na classificação facetada e taxonomias dinâmicas. 2013, 233f. Tese (Doutorado em Ciência da Informação). Curso de Pós Graduação em Ciência da Informação – Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Minas Gerais, 2013. Disponível em: <<http://www.bibliotecadigital.ufmg.br/dspace/handle/1843/BUOS-97FGQ6>>. Acesso em: 30 jul. 2014.

RANGANATHAN, S. R. **Prolegomena to library classification**. Bombay: Asia Publ. House, 1967.

SARACEVIC, T. Ciência da Informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**, v. 1, n. 1, p.41-62, 1996. Disponível em: <portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/235/22>. Acesso em: 29 jul. 2014.

SILVA, E. L.; MENEZES, E. M. **Metodologia da pesquisa e elaboração de dissertação**. 4. ed. Florianópolis: UFSC, 2005. 138 p. Disponível em: <<http://www.praticadapesquisa.com.br/2011/08/metodologia-da-pesquisa-e-elaboracao-de.html>>. Acesso em: 24 abr. 2014.

VICKERY, B. C. **Classificação e indexação nas ciências**. Rio de Janeiro: BNG/Brasilart, 1980.

Modalidade da apresentação: Pôster

APLICAÇÃO DE MODELO DE REENGENHARIA EM TESAURO

APPLICATION OF REENGINEERING MODEL ON THESAURUS

Benildes Coura Moreira dos Santos Maculan
Gercina Ângela Borem de Oliveira Lima

Resumo: Pesquisa de doutorado em andamento, com o objetivo de estudar um modelo de reengenharia de sistemas de organização do conhecimento (SOCs) para transformar um tesauro tradicional em um SOC, convergente com o meio digital. A modelagem será aplicada no tesauro brasileiro THESAGRO, dentro do recorte temático da “intensificação agropecuária”. Como fundamento teórico-contextual apresentam-se conteúdos sobre semântica, *web* semântica, os SOCs e seus relacionamentos semânticos, o conceito e a definição que o delimita, e sobre a teoria das valências dos verbos. A metodologia será aplicada a partir do modelo proposto por Soergel *et al.* (2004), que consiste no aperfeiçoamento e na explicitação formal dos relacionamentos semânticos na estrutura conceitual de um tesauro tradicional. O *software* para a implantação do modelo será o E-Termos, que é um sistema colaborativo para gestão terminológica. Espera-se validar o modelo aplicado, demonstrando a viabilidade da conversão de um tesauro em uma estrutura formalizada, possibilitando obter interoperabilidade entre os distintos SOCs e entre diferentes sistemas de recuperação de informação (SRI).

Palavras-chave: Reengenharia de tesauro. Relacionamentos semânticos. Sistemas de recuperação de informação. Sistema de organização do conhecimento.

Abstract: Doctoral research in progress aiming at studying a reengineering model to convert a traditional thesaurus in a knowledge organization system (KOS), according to the digital environment. The modeling will be applied to the Brazilian THESAGRO thesaurus and the extract study is the sub-area “agricultural intensification”. The theoretical and contextual basis of this research encompass themes such as semantics, semantic web, KOS and their conceptual relationships, the concept and definition, which defines each concept, and the theory of valences of verbs. The methodology will depart from the model proposed by Soergel *et al.* (2004), which consists in the improvement and formal expression of the semantic relationships into the conceptual structure of a traditional thesaurus. The software for the implementation of the model will be the E-terms, which is a collaborative system for terminology management. We expect to validate the applied model, demonstrating the feasibility of the conversion of a thesaurus in a formalized structure, thus possibly enabling interoperability among different KOS and between different information retrieval systems.

Keywords: Thesaurus reengineering. Semantic relationships. Information retrieval systems. Knowledge organization system.

1 INTRODUÇÃO

O termo “sistema de organização do conhecimento” (SOC) é relativamente novo e é usado para nomear os instrumentos tais como os sistemas de classificação, lista de cabeçalho de assuntos, tesouros, ontologias, entre outros. Segundo Bräscher (2009), os SOC são sistemas conceituais criados para sistematizar os conceitos e suas relações, dentro de um domínio, e para

auxiliar a recuperação de informações. Eles têm como objetivo controlar os processos de representação, classificação, ordenação e armazenamento dos recursos informacionais, visando a recuperação e comunicação de informações.

O SOC do tipo tesouro é composto por um conjunto de termos relacionados semanticamente e genericamente entre si, permitindo diferentes tipos de organização (ex.: alfabética, relacional, estruturada por campos semânticos, entre outros). Apesar de possuírem uma estrutura semântica, eles ainda não são construídos de forma adequada para uso em ambiente digital. A estrutura de um tesouro geralmente é constituída por relacionamentos entre conceitos, que abrangem relações de equivalência, hierárquicas e associativas. Essa estrutura conceitual vem evoluindo (MOTTA, 1987; CAMPOS, 1995; CAMPOS; GOMES, 2003; CAMPOS, 2004, entre outros), e já há evidências de diferentes desdobramentos para cada um desses tipos de relações. Porém, os relacionamentos explicitados na estrutura do tesouro ainda não determinam a diversidade dos tipos existentes. Podemos citar, por exemplo, o caso da relação de equivalência, pois ela pode variar dependendo de uma equivalência ortográfica, total ou parcial, de uso de abreviaturas, nome fantasia ou equivalência em outro idioma. Essa situação se repete, também, para os relacionamentos hierárquicos e associativos.

Nesse contexto, o objetivo desta pesquisa é aplicar um modelo de reengenharia de tesouro tradicional para torná-lo um instrumento mais formalizado, de tal forma que os relacionamentos semânticos entre os conceitos do sistema estejam identificados e explicitados para o usuário. O foco recairá, sobretudo, no apontamento dos relacionamentos associativos.

2 FUNDAMENTOS TEÓRICOS

Na área da Organização do Conhecimento, Svenonius (2000) discorreu sobre a estrutura semântica usada na normalização de vocabulários controlados, que é conseguida através do estabelecimento de relacionamentos entre os termos e seus referentes. A autora recomenda que a desambiguação terminológica não deva ser realizada apenas no nível da determinação do domínio, pois isso pode limitar a compatibilidade entre distintos vocabulários.

Existem diferentes semânticas para a desambiguação terminológica. Entre elas está a semântica linguística, que é uma disciplina da Linguística e estuda o significado das expressões (fonemas, morfemas, palavras, sintagmas, frases). Ela também analisa as relações de significado que essas expressões estabelecem entre si e o domínio ou o mundo. Concordamos com Katz (1982), que defende a ideia de que qualquer disciplina que se proponha a estudar significados terminológicos precisa postular sobre fenômenos semânticos:

sinonímia e paráfrase; antonímia; hiperonímia/hiponímia; ambiguidade; redundância semântica; contradição; sinteticidade; fenômeno das informações implícitas (implicação ou pressuposição). Assim, ao trabalharmos no campo da Ciência da Informação na criação de instrumentos terminológicos e conceituais, é preciso buscar respaldo na Linguística para entender os fenômenos semânticos.

Nos estudos sobre semântica, Santos e Cardoso (2007) destacam três grandes escolas de análise: (1) denotacionista, (2) funcionalista e (3) pragmática. Na primeira, os símbolos substituem os objetos, pois estão unidos aos seus respectivos significantes. Na segunda, os símbolos representam uma relação, dentro da própria língua, entre diferentes objetos, no qual o significado de uma palavra ou sintagma¹⁷⁹ tem estreita relação com a função que desempenha dentro do texto. Já na terceira escola, da pragmática, o sentido é definido a partir do contexto. Isso é, leva-se em consideração a função significativa do símbolo, analisado em relação ao contexto real de utilização desse símbolo (SANTOS; CARDOSO, 2007).

Ainda no campo da semântica, há duas distintas unidades de análise: (1) a semântica da palavra, que foca nas relações (sinonímia, antonímia, hiperonímia, hiponímia, ambiguidade lexical, polissemia, homonímia, metáfora e metonímia) de significado entre palavras ou entre morfemas¹⁸⁰; (2) a semântica da frase, que prioriza a ambiguidade estrutural, as anomalias no arranjo das palavras na frase, as relações de paráfrase e contradição, a implicação semântica e a pressuposição (PEDROSA, 2001). Pedrosa (2001) acrescenta o plano da semântica do verbo, e aponta que há restrição na seleção semântica dos verbos.

Os verbos carregam um valor semântico e compreender esse valor é importante na determinação de relacionamentos nos tesouros (SOERGEL *et al.*, 2004). No sentido de melhor apreender esse assunto, buscamos respaldo na “Abordagem Sintático-Semântica da Oração com base na Estrutura Argumental”, também denominada de “Gramática de Valências”¹⁸¹ ou “Gramática de Casos”, proveniente da área da Linguística.

¹⁷⁹ Sintagma é um segmento linguístico existente em qualquer enunciado, onde os signos ligam-se uns aos outros formando grupos, expressa em uma relação de dependência a partir de regras sintagmáticas (sequência correta de uma frase). Nessa relação há um elemento determinado e outro determinante (ou subordinado), e cada um deles constitui um sintagma. Os sintagmas nominal e verbal são os mais básicos, e ainda existem os sintagmas adjetival, adverbial e preposicional. Ex.: As margaridas (sintagma nominal) murcharam (sintagma verbal).

¹⁸⁰ Morfema é a menor parte significativa de uma palavra, quando a analisamos morfológicamente. Por exemplo, a palavra “sol” possui apenas um morfema; mas a palavra “insolação” terá três morfemas: in(prefixo)+sol(radical)+ação(sufixo).

¹⁸¹ Neste artigo decidimos por usar o termo “Gramática de Valências” para denominar essa abordagem.

Borba (1996, p. xxi) conceitua “valência verbal” como o “conjunto de relações estabelecidas entre o verbo e seus argumentos ou constituintes indispensáveis”. O autor afirma que os verbos assinalam tipos de situação, pois, em geral, uma frase designa uma situação. Para determinar a situação à qual um verbo pertence é necessário saber qual a natureza das entidades designadas por ele. Marques (2003) aponta que na Gramática de Valências o verbo é o elemento central e que sua relação com os demais elementos da frase são estudados sob o ponto de vista sintático e semântico. Dessa forma, os diversificados aspectos semânticos dos verbos podem auxiliar na determinação dos seus papéis temáticos de forma mais rigorosa, tendo um regulamento teórico de uma teoria gramatical. Esse tópico é de interesse deste estudo, uma vez que os relacionamentos dentro do tesouro serão ancorados na determinação das expressões verbais, que mantêm o sistema de conceitos consistente.

A questão do conceito é estudada a partir do pensamento e estudos de autores tais como Dahlberg (1978a), Ranganathan (1967), Bakhtin (1995) e Vygotsky (1998). Dahlberg (1978a) indica o “conceito” como “unidade do conhecimento”, e esse conhecimento é constituído pelo conjunto de enunciados verdadeiros sobre um objeto e representado por uma forma verbal, que é o termo. Por outro lado, Ranganathan (1967) apresenta o “conceito” como uma “unidade do pensamento”. Por seu turno, Bakhtin (1995) aponta que não pode haver comunicação verbal desligada de uma situação concreta, acrescentando que os enunciados ou conceitos produzidos estão sempre em um contexto ou domínio social, no qual são organizados. Já Vygotsky (1998) pondera que quando se trata do pensamento verbal, os conceitos não são inatos, pois são construções culturais, criados como consequência da socialização, ou seja, são contextualizados.

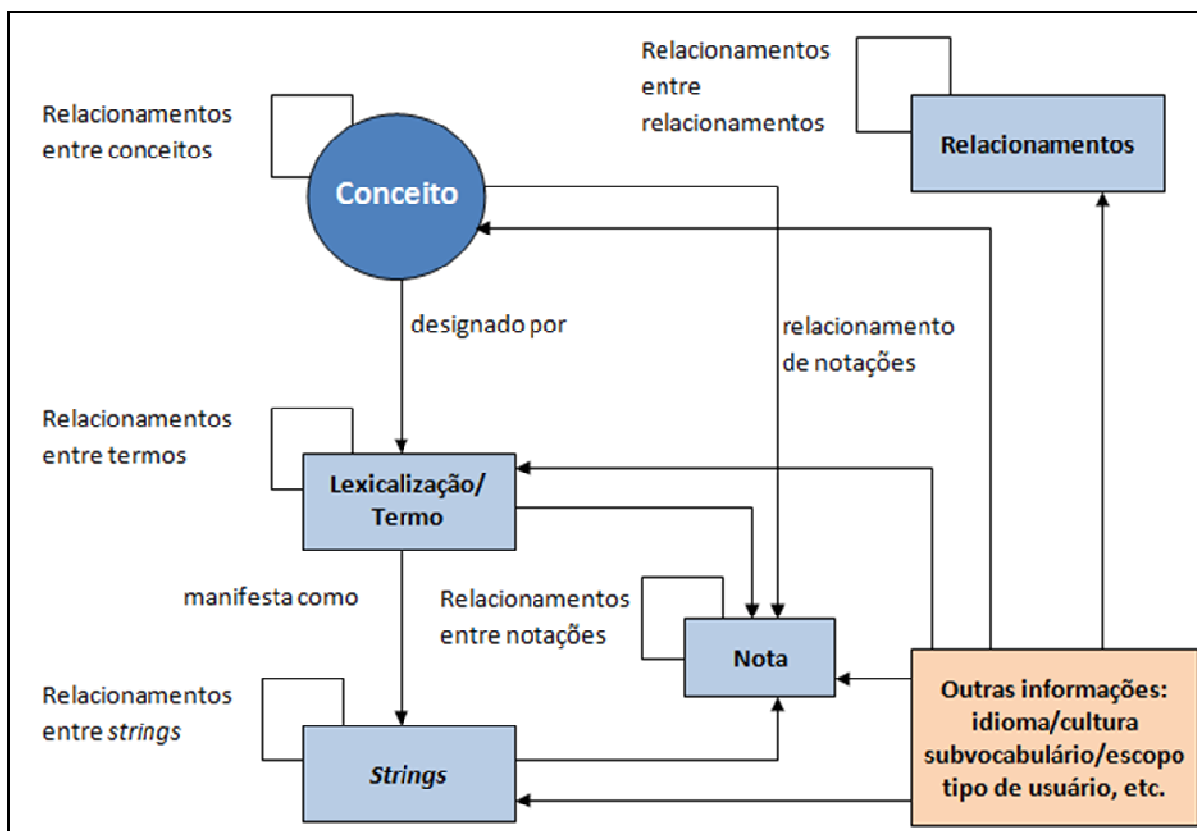
Na contextualização de um conceito, a definição tem um importante papel, pois permite a sua identificação e delimitação em determinado domínio. Nesse sentido, Dahlberg (1978b; 1992) apresenta a função da definição no estabelecimento de uma equivalência entre o termo-características-referente, com o propósito de delimitar o uso, em dado domínio, de um dado conceito. Essa definição é essencial na determinação das relações semânticas existentes na estrutura conceitual do tesouro, no qual os relacionamentos compõem a sua rede paradigmática e sintagmática. A rede paradigmática é constituída por relações lógico-semânticas para organização dos descritores (*a priori*), tais como as relações gênero/espécie, todo/parte e de oposição (homem/mulher). Já a rede sintagmática é formada pelas relações entre termos-descritores, validadas no contexto (*a posteriori*), numa relação funcional, que pode ocorrer entre elementos compostos ou complexos de um tesouro, para a descrição do conteúdo dos itens de informação. Os aportes sobre os relacionamentos advêm de Foskett

(1973), Dahlberg (1978b), Farradane (1980), Gomes (1990), Svenonius (2000), Campos (2001), Green (2001), Moreira e Lara (2011), entre outros.

2 METODOLOGIA

Soergel *et al.* (2004) propõem uma modelagem conceitual de reengenharia cuja principal característica é a clara separação nos níveis de conceito, de termo ou lexicalização e de *string*. A estrutura conceitual do modelo permite que as informações formalizadas sobre os distintos níveis sejam individualizadas, podendo ocorrer relacionamentos entre entidades do mesmo tipo (por exemplo, entre conceitos, entre *strings* ou entre relacionamentos), como também podem ser estabelecidas relações entre as diferentes entidades, conforme pode ser observado na FIG.1.

FIGURA 1 – Modelo conceitual para transição de SOC tradicional em SOC avançado



Fonte: traduzido de Soergel *et al.* (2004, p.11).

Este modelo será aplicado no tesouro THESAGRO brasileiro, com um recorte na temática "intensificação agropecuária". O THESAGRO é mantido pela Biblioteca Nacional de Agricultura (BINAGRI), órgão da Secretaria de Executiva do Ministério da Agricultura, Pecuária e Abastecimento.

Os procedimentos na aplicação do modelo de Soergel *et al.* (2004) serão os seguintes:

Etapa 1: Definir a estrutura do SOC avançado

Nesta etapa, inicia-se com o mapeamento da estrutura conceitual do domínio, e, em seguida, determina-se o conjunto de tipos de relacionamentos que irá compor essa estrutura conceitual, a partir da identificação dos relacionamentos existentes entre as diferentes entidades. A partir do modelo, será possível identificar relacionamento entre os distintos elementos, que são tratados de forma independente: (1) conceitos; (2) termos; (3) *strings*; (4) notas; (5) relacionamentos.

Etapa 2: Coletar a terminologia e outras informações de um ou mais SOCs

Em primeiro lugar, será mapeado o conjunto de conceitos e termos que representam o domínio escolhido, ou seja, da “intensificação agropecuária”. Para esta etapa os autores recomendam que a terminologia do domínio deve ser enriquecida com a agregação de informações adquiridas de um ou mais instrumentos (sistemas de classificação, tesouros, entre outros), e na literatura (divulgação científica, índices de publicações periódicas e de resumos, teses e dissertações) de áreas correlatas. Assim, haverá a coleta de terminologia e outras informações em diferentes instrumentos, dentre os quais já foram selecionados: (1) AGROVOC, um tesouro multilíngue criado pela Organização das Nações Unidas para a Alimentação e Agricultura (FAO) e pela Comissão das Comunidades Europeias, e abarca todos os campos temáticos da agricultura, silvicultura, pesca, seguridade alimentária e domínios associados; (2) NAL THESAURUS, um tesouro produzido em colaboração entre a Biblioteca Nacional de Agricultura, e o Instituto Interamericano de Cooperação para a Agricultura, que abrange termos da agricultura para indexação do banco de dados bibliográficos; (3) THESBIO: tesouro estruturado por disciplinas e conceitos correlatos às ciências da vida, especializado em biodiversidade brasileira, com mais de 5.000 termos.

Etapa 3: Aplicar a reengenharia do tesouro, por meio de um *software*, editor de tesouro

- a. tornar a informação existente mais precisa;
- b. adicionar novas informações.

Esta etapa é considerada por Soergel *et al.* (2004) a etapa mais trabalhosa do modelo, que é usar um *software* para edição manual ou semi-automática do tesouro. Isso será realizado utilizando o *software* e-Termos¹⁸² (Termos Eletrônicos) que é um ambiente (*web*), colaborativo para gestão terminológica, desenvolvido no âmbito de uma parceria entre a Embrapa Informática Agropecuária (CNPTIA), Universidade de São Paulo (USP Campus de

¹⁸² Disponível em: <<https://www.etermos.cnptia.embrapa.br/index.php>>.

São Carlos, SP) e Universidade Federal de São Carlos (UFSCar). Ele foi criado com base em pressupostos teóricos da linguística e possui diferentes ferramentas de análise linguística que dão suporte às tarefas de Processamento de Língua Natural (PLN).

3 RESULTADOS PARCIAIS E CONSIDERAÇÕES FINAIS

Alguns resultados já podem ser observados, tais como a compilação de um *corpus* de termos descritores no domínio temático da intensificação agropecuária, assim como a definição para tais termos. A partir dessas definições, está sendo possível determinar os relacionamentos existentes entre conceitos, termos e *strings*. Com este estudo, espera-se validar o modelo aplicado, demonstrando a viabilidade da conversão de um tesauro tradicional em uma estrutura formalizada, com o intuito de obter interoperabilidade entre distintos vocabulários e entre diferentes sistemas de recuperação de informação (SRI).

AGRADECIMENTOS

Agradecemos apoio financeiro concedido para esta pesquisa e para o Grupo de Pesquisa MHTX através de financiamento ou chamadas por editais pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela Fundação de Amparo à Pesquisa do estado de Minas Gerais (FAPEMIG) e pela Pró-Reitoria de Pesquisa da UFMG (PRPq).

REFERÊNCIAS

BAKHTIN, Mikhail. **Marxismo e filosofia da linguagem**. 7. ed. São Paulo: Hucitec, 1995.

BORBA, Francisco da S. **Uma gramática de valências para o português**. São Paulo: Ática, 1996.

BRÄSCHER, 2009. Tesauro, taxonomia e ontologia: uma evolução?. In: CICLO DE CONFERÊNCIAS SOBRE ORGANIZAÇÃO DA INFORMAÇÃO LEGISLATIVA E JURÍDICA. **Palestra**. Brasília, 2009. Disponível em: <<http://projeto.lexml.gov.br/arqs/BRASCHER.pdf>>. Acesso em: 19 jul. 2014.

CAMPOS, Maria Luiza de A. **Linguagem documentária: teorias que fundamentam sua elaboração**. Rio de Janeiro: EUFF, 2001.

CAMPOS, Maria Luiza de A. Linguagens documentárias: núcleo básico de conhecimento para seu estudo. **Revista da Escola de Biblioteconomia UFMG**, Belo Horizonte, v. 24, n. 1, p. 52-62, jan./jun. 1995.

CAMPOS, Maria Luíza de A. Modelização de domínios de conhecimento: uma investigação de princípios fundamentais. **Ciência da Informação**, Brasília, v. 33, n. 1, p. 22-32, jan./abr. 2004.

CAMPOS, Maria Luiza de A.; GOMES, Hagar Espanha. Organização de domínios de conhecimento e os princípios ranganathianos. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 8, n. 2, jul./dez. 2003.

DAHLBERG, Ingetraud. Knowledge organization and terminology; philosophical and linguistic bases. **International Classification**, v. 19, n. 2, p. 65-71, 1992.

DAHLBERG, Ingetraud. A referent-oriented, analytical concept theory of Interconcept. **International Classification**, v. 5, n. 3, p. 122-151, 1978b.

DAHLBERG, Ingetraud. Teoria do conceito. **Ciência da Informação**, Brasília, v. 7, n. 2, p. 101-07, 1978a.

FARRADANE, J. Relational Indexing: part I e part II. **Journal of Information Science**, n.1, p. 267-276 e 313-324, 1980. Disponível em: <http://www.fims.uwo.ca/people/faculty/Frohmann/LIS677/Documents/Subject%20Analysis/Farradane_1.pdf>. Acesso em: 12 jun. 2014.

FOSKETT, A.C. **A abordagem temática da informação**. São Paulo: Polígono, 1973.

GOMES, Hagar Espanha (Org.). **Manual de elaboração de tesouros monolíngues**. Brasília: Programa Nacional de Bibliotecas de Instituições de Ensino Superior, 1990.

GREEN, Rebecca. Overview of relationship in knowledge organization. In: BEAN, Carol A.; GREEN, Rebecca (Ed.). **Relationship in knowledge organization**. Dordrecht: Kluwer, 2001. Chapter 1, p. 3-18.

KATZ, Jerrold J. O escopo da semântica. In: DASCAL, Marcelo (Org.). **Fundamentos metodológicos da linguística: semântica**. Campinas: UNICAMP, 1982. v. 3, p. 43-61.

MARQUES, Ana Lucia de Andrade. **As relações entre verbos de movimento e complementos locativos de origem e meta com enfoque da gramática de valências**. 408f. Tese (Doutorado) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2003.

MOREIRA, Walter; LARA, Marilda L. G. de. Relações conceituais e categorias filosóficas: aportes das ontologias e da terminologia para a representação do conhecimento. In: XII Encontro Nacional de Pesquisa em Ciência da Informação, 12., 2011, Brasília, DF. **Anais...** Brasília: ENANCIB, 2011.

MOTTA, Dilza F. da. **Método relacional como nova abordagem para a construção de tesouros**. 89f. 1987. Dissertação (Mestrado em Ciência da Informação), Instituto Brasileiro de Informação em Ciência e Tecnologia. Rio de Janeiro: IBICT, 1987. Disponível em: <<http://www.conexaorio.com/bitl/dilza/>>. Acesso em: 12 jun. 2014.

PEDROSA, Cleide Emília F. Recursos para uma desambiguação das “frases” veiculadas pelas revistas Veja e Isto É. In: V CONGRESSO NACIONAL DE LINGUÍSTICA E FILOLOGIA, 5., 27 a 31 de agosto de 2001, Rio de Janeiro. **Anais...** Rio de Janeiro: UERJ, Cadernos do CNLF, v. 5, n. 11, 2001.

RANGANATHAN, S. R. **Prolegomena to library classification**. Bombay: Asia Publishing House, 1967.

SANTOS, Diana; CARDOSO, Nuno (Ed.). **Reconhecimento de entidades mencionadas em português: documentação e actas do HAREM**: a primeira avaliação conjunta na área. [S.l.]: Repositório Linguatca, 2007. cap. 4, p. 43–57.

SOERGEL, Dagobert *et al.* Reengineering thesauri for new applications: the Agrovoc example. **Journal of Digital Information**, v. 4, n. 4, 2004.

SVENONIUS, E. The intellectual foundations of information organization. Cambridge: The MIT Press, 2000.

VYGOTSKY, Lev S. **Pensamento e linguagem**. 2. ed. São Paulo: Martins Fontes, 1998.

CIÊNCIA DA INFORMAÇÃO, POSITIVISMO E MEMÓRIA: RELAÇÕES COM A REDOCUMENTARIZAÇÃO DO DOCUMENTO DE VALOR HISTÓRICO

INFORMATION SCIENCE, POSITIVISM AND MEMORY: RELATIONS WITH REDOCUMENTARIZATION OF THE DOCUMENT OF HISTORICAL VALUE

Hercules Pimenta dos Santos

Resumo: A pesquisa de doutorado em andamento versa sobre um estudo teórico interdisciplinar nas práticas e processos que tratam dos procedimentos de Redocumentarização para se compreender os seus impactos no fazer historiográfico atual. Busca-se refletir sobre a noção polissêmica de documento para as áreas de Ciência da Informação e História, com base em conceitos capazes de elucidar a interpretação de uma realidade específica, constituída por discussões muitas das vezes carregadas de aspectos subjetivos. Acredita-se que a partir desse debate será possível lançar novas perspectivas teóricas sobre a importância desse objeto em ambas as áreas do conhecimento. O ponto de partida são as necessidades informacionais produzidas na contemporaneidade, procurando questionar o documento de arquivo digitalizado e refletir sobre os seus limites, desejando contribuir para alargar a compreensão desta amplitude que se apresenta como um problema de ordem teórico-prática dentro da viabilização do estudo em desenvolvimento. Se conclui, até o presente momento, que o documento ao ser redocumentarizado, de forma a manter um forte vínculo deste com a tradição positivista, dificulta as abordagens teóricas e práticas concebidas para o fazer historiográfico atual, por se considerar haver impactos negativos se ao historiador não restar muito que interpretar, além do que for objetivado pelo produtor do documento, ao definir, de antemão, o que interessaria ao consulente.

Palavras-chave: Ciência da Informação, Documento digitalizado, Redocumentarização, Memória, Positivism.

Abstract: The doctorate level research, in progress, is an interdisciplinary theoretical study on the practices and processes that deal with procedures of Redocumentarization, to understand their impact on the current historiographical methods. Seeks to reflect on the polysemic notion of document for the areas of Information Science and History, based on concepts that able us to elucidate a specific interpretation of reality, consisting of discussions often laden with subjective aspects. It is believed that this discussion will allow for new theoretical perspectives on the importance of this object in both areas of knowledge. The starting point are the informational needs produced in contemporary times, seeking to question the scanned document and reflect on their limits, wanting to contribute to broaden understanding of this magnitude that is presented as a problem of theoretical and practical study into the developing of this research. Concludes, so far, that the document after the Redocumentarization, maintaining a strong bond with this positivist tradition, hampers, theoretical and practice, to make the current historiographical approaches, by considering whether there negative impacts to historian if not much is left to interpret beyond what is objectified by the producer of the document, if his define, in advance, what interest the inquirer.

Keywords: Information Science, Scanned Document, Redocumentarization, Memory, Positivism.

1 INTRODUÇÃO

Este texto tem por objetivo apresentar resultados iniciais e relações teóricas com a pesquisa de doutorado “Redocumentarização do Documento de Valor Histórico: Impactos da Inovação Tecnológica na Tradição Historiográfica”. A pesquisa versa sobre um estudo teórico interdisciplinar nas práticas e processos que tratam dos procedimentos de Redocumentarização¹⁸³ a partir da perspectiva da Organização da Informação. Busca-se compreender os seus impactos no fazer historiográfico atual, tendo como base teórica, fundamentando tal questionamento, a noção ampliada do documento de valor histórico, preconizada pelo Movimento dos *Annales*¹⁸⁴.

Os desafios enfrentados pela Organização da Informação são muito anteriores ao desenvolvimento das tecnologias atuais, gerando o entendimento do relevante estudo dos conceitos tradicionalmente usados no contexto das unidades de informação, buscando contextualizar suas mudanças, além de possíveis deficiências e melhorias nos processos atualmente adotados.

2 RELAÇÕES: CIÊNCIA DA INFORMAÇÃO, POSITIVISMO E MEMÓRIA

Procura-se buscar na área da Ciência da Informação noções tradicionalmente usadas para caracterizar os atores e os objetos envolvidos nos processos ligados à informação, tais como o documento, o ciclo documental e a organização da informação, objetivando identificar impactos no fazer historiográfico, a partir da migração dos documentos de valor histórico para o ambiente digital. Um aspecto importante para esta relação interdisciplinar decorre das demandas sociais relativas à preservação e à atualização da Memória coletiva em sua relação com o estudo da cultura material¹⁸⁵.

É notório que o conceito de documento, nos últimos vinte anos, vem constituindo um novo centro de debates como objeto da Ciência da Informação. O movimento também nomeado de "neodocumentação" ou "redocumentalização" apresenta o deslocamento da informação como objeto de estudo daquela ciência para o documento, constituindo a

¹⁸³ Redocumentarizar significa tratar um documento, ou coleções documentais, rearticulando os conteúdos. Trata-se de uma série de procedimentos que objetivam transpor o documento tradicional para o meio digital.

¹⁸⁴ O movimento dos *Annales*, em parte, surge da crítica à instrumentalização das fontes documentais oficiais, ou de arquivo oficial, que perpassou do fim do século XIX ao início do XX, momento mais expressivo das abordagens por meio dos estudos filosóficos positivistas.

¹⁸⁵ O documento se destacou por ser um espaço privilegiado da memória, em particular, por fazer parte dos objetos que ilustram a cultura material (MENESES, 1998; MURGUIA, 2008) e a memória cultural (DODEBEI, 1997) e por se relacionar com o sistema informativo documental (RENDÓN ROJAS, 2005; 2008).

"materialidade da informação" (FROHMANN, 2008). Isto alimenta os estudos das práticas sociais da informação com suas consequências e efeitos, além de uma maior proximidade entre diversas áreas de preservação da Memória que lidam com o documento (LARA, 2010).

2.1 - Problema de pesquisa

O fazer historiográfico no começo do século XX passou a ser baseado em uma noção documental ampliada, ao se questionar os tipos de informação intrínseca que estes artefatos culturais podem conter, especialmente, de conteúdo histórico. A presente investigação está buscando indagar a manutenção destas possibilidades nos processos de Redocumentarização. A preocupação está na necessária presença dos atributos intrínsecos da natureza físico química do objeto, isto é, poder ter acesso à sua forma geométrica, cor, dureza e odor, entre outras fundamentais características, mesmo após a digitalização e virtualização da documentação.

O processo de Redocumentarização é considerado uma imposição do universo digital, e versaria sobre tratar novamente documentos ou coleções, rearticulando os conteúdos segundo o uso ou a interpretação do usuário (ARNAUD; MERZEAU, 2009). O seu ciclo passa por uma série de procedimentos, objetivando transpor o documento tradicional para o meio digital (SALAÜN, 2009).

O processo de redocumentarizar, como tratamento digital de livros e coleções é objeto de projetos que buscam a promoção do acesso e uso público da Memória do saber produzido historicamente e de interesse da pesquisa científica, descentralizando o acesso à informação histórica de caráter público.

Considera-se este processo como outra ordem documentária que é definida a partir de sua relação com os dispositivos existentes em diferentes sociedades e épocas. Estas proposições acarretam em algo que nesta pesquisa será primordial refletir sobre, que é melhor entender o que estes autores consideram por rearticular os conteúdos segundo o uso ou a interpretação do usuário (ARNAUD; MERZEAU, 2009). Pois, historiograficamente, o documento escrito foi considerado um instrumento secundário de informação. Assim, ao ser redocumentarizado, de forma a manter um forte vínculo deste com a tradição positivista, isto dificultaria as abordagens teóricas e práticas concebidas para o fazer historiográfico atual.

O conceito de documento ora focalizado está baseado em um contexto de trabalho pouco influenciado por artefatos tecnológicos. Está-se considerando que as mudanças estruturais no processo informacional causadas pela digitalização dos documentos e pela comunicação eletrônica trouxeram impactos definitivos na configuração dos documentos.

2.2 Concepções da noção de documento para a História e para a Ciência da Informação: influência positivista

No século XIX temos o surgimento das tentativas de se explicar e de se interpretar os fenômenos sociais, alegado pela necessidade de se relacionar a filosofia e as ciências naturais como forma de se obter o entendimento dos processos humanos. Para tanto, buscou-se bases científicas em métodos de observação, experimentação, dedução, indução e comparação. Um dos expoentes mais significativos desse pensamento foi o movimento filosófico criado pelo francês Auguste Comte, o Positivismo (COMTE, 1972). Além de ter cunhado o termo, Comte foi precursor no demarcar a sociedade como objeto de observação. Hoje, a palavra positivismo é muitas vezes apropriada como sinônimo de empirismo, pensada como um ponto de vista que negligencia o trabalho teórico dos pesquisadores. Segundo Hjørland (2005) isto não está, no entanto, em conformidade com o positivismo clássico.

Seguindo a linha de pensamento iniciada por Auguste Comte, um dos desdobramentos da História Científica foi a determinação da dita “História Positivista” ou Metódica, a qual teve como um de seus mais expressivos precursores o historiador alemão Leopold von Ranke. A dita “História Positivista”, de Ranke, tinha como princípio escrever a história como realmente ela teria acontecido buscando o “apelo à objetividade” (GAY, 1990, p. 70). Ranke propôs que a narrativa e o registro histórico fossem provenientes da pesquisa sobre fontes documentais primárias, ou seja, dos documentos de arquivo. Assim, Ranke colaborou para que o arquivo passasse a ser o lugar por excelência do historiador. Desta forma, a contribuição de Ranke à historiografia está “em sua visão enaltecida dos documentos” (Gay, 1990, p. 74).

No entanto, sabemos que considerar somente o método da crítica textual de documentos de arquivo leva a algumas limitações conceituais de documento histórico, no sentido de que a fonte histórica seria o texto, ou documento escrito, para a geração de outro texto. A apropriação do documento como fonte histórica fora reconhecida academicamente e explicitada com a noção de História Metódica, ou dita Positivista, de Ranke. Esta corrente possui harmonia com o chamado “espírito positivo”, proveniente das ideias de Auguste Comte. Ranke propôs o aperfeiçoamento do método por meio da consulta às fontes documentais e o desenvolvimento de uma teoria histórica.

No universo historiográfico passamos a contar com diversos termos. Dentre estes, o uso de “História Positivista” que para Chartier (1978) obteve maior aceitabilidade, porém denotando sentido negativo por se referir à forma que “não convém para fazer história”. Tal situação acabou por depreciar a filosofia positivista proveniente de Auguste Comte.

A partir da “revolução historiográfica”, proporcionada pelo movimento dos *Annales*, temos o início dos questionamentos ao método e a abordagem da dita “História Positivista” considerada metodologicamente limitada. Passa-se a entender que, para se obter entendimentos satisfatórios sobre esses efeitos, o historiador necessitaria realizar mediações de outras ordens (BURKE, 1992).

Abordando a perspectiva informacional, Smit & Guimarães (2001) consideram que os documentos são fruto de um conjunto de procedimentos objetivando expressar o conteúdo destes para facilitar a recuperação da informação. Enquanto um produto documental decorre de uma etapa analítica propriamente dita, ou seja, uma “análise dentro da análise” (GUIMARÃES, 2003); o outro advém de uma etapa ulterior de síntese. Aqui, percebemos que o seu objetivo é propiciar precisão no processo de recuperação da informação. Assim, entendemos que um valor maior atribuído aos aspectos objetivos do processo de análise, poderia desconsiderar os aspectos materiais, subjetivos e sociais. Para refletir sobre isto nos apropriamos de Rendón Rojas (2005) que considera este ato tratar-se do documento elaborado e/ou de interesse de um especialista que entra na jurisdição do profissional da informação documental. Sujeito este, responsável por inserir o documento em um determinado sistema informativo de modo a institucionalizá-lo, agregando-lhe informação ao representá-lo, ao organizá-lo e ao torná-lo acessível ao usuário. As ações deste profissional da informação documental atribui novo sentido ao documento inicial, pois o coloca dentro de um contexto específico diferente do de sua proveniência podendo ainda eliminar características importantes.

3 CONSIDERAÇÕES FINAIS

Os historiadores dos *Annales* criticaram o método de análise das fontes que se baseava apenas na superficialidade fenomênica expressa nos documentos oficiais escritos. Importa ressaltar que, o problema não estaria na utilização dessas fontes, mas na compreensão de que elas seriam as únicas válidas para a escrita da História, assim valorizadas como as únicas que devem ser preservadas. Entende-se aqui que tais procedimentos se aproximam daquilo que se constituía como uma concepção positivista “ingênua” de documento, ao privilegiar a organização deste tipo de fonte de informação (a de caráter oficial) apenas a partir do seu conteúdo informacional e ao lhe atribuir um novo contexto sem, por exemplo, agregar à sua descrição a proveniência e/ou características físicas originais. Diante do contexto da Redocumentarização a partir da perspectiva da Organização da Informação, a reflexão que surge é que seriam procedimentos adotados que impactam as possibilidades de interpretação

historiográfica. No desenvolvimento da pesquisa em foco, está-se refletindo sobre o quanto e como ocorreria este conflito, além de procurar relacionar o que pode ser feito para amenizar o que possa se constituir em impactos negativos.

REFERÊNCIAS BIBLIOGRÁFICAS

ARNAUD, M; MERZEAU, L. (Org.). Traçabilité et Réseaux Hermès. **CNRS Éditions**, Paris, n.53, avril 2009.

BURKE, P. (Org.). **A escrita da história: novas perspectivas**. Trad. Magda Lopes. São Paulo: UNESP, 1992.

CHARTIER, R. História positivista. In: LE GOFF, J. *et al.* (Org.). **A nova história**. Coimbra: Almedina, 1978.

COMTE, A. **Opúsculos de Filosofia Social: 1819-1828**. Porto Alegre; São Paulo: Globo/USP, 1972.

DODEBEI, V. L. D. **O sentido e o significado de documento para a memória social**. Tese (Doutorado em Comunicação) – Escola de Comunicação, UFRJ, Rio de Janeiro, 1997.

FROHMANN, B. O caráter social, material e público da informação. In: FUJITA, M.S.L., MARTELETO, R.M., LARA, M.L.G. (Org.). **A dimensão epistemológica da Ciência da Informação e suas interfaces, políticas e institucionais nos processos de produção, acesso e disseminação da informação**. SP: Cultura Acadêmica; Marília; Fundepe, 2008.

GAY, Peter. **O Estilo na História**. São Paulo: Companhia das Letras, 1990.

GUIMARÃES, J. A. C. A análise documentária no âmbito do tratamento da informação: elementos históricos e conceituais. In: RODRIGUES, J.M.; LOPES, I. L. (Org.) **Organização e representação do conhecimento na perspectiva da ciência da informação**. Brasília: Thesaurus, 2003.

GUIMARÃES, J. A. C.; RABELLO, R. A contribuição metodológica da diplomática para a análise documental de conteúdo em arquivos e bibliotecas. In: RICHTER, E. I. S.; ARAUJO, J. C. G. (Org.). **Paleografia e Diplomática no curso de Arquivologia**. Santa Maria: FACOS-UFSM, 2007.

HJØRLAND, B. Empiricism, rationalism and positivism in library and information science. **Journal of Documentation**, v. 61, n. 1, 2005.

LARA, M. L. G. Documento e significação na trajetória epistemológica da Ciência da Informação. IN: FREITAS, L. S. (Org.) **Documento: Gênese e contextos de uso**. Niterói: Ed. UFF, 2010.

LE GOFF, J. **História e memória**. 3. ed. Campinas: UNICAMP, 1994. (Coleção Repertórios).

MENESES, U. T. B. Memória e cultura material: documentos materiais no espaço público. **Estudos Históricos**, Rio de Janeiro, v. 11, n. 21, 1998.

MURGUIA, E. I. A produção social do documento: valor, informação e instituição. Um lugar de reflexão para as Ciências Humanas. In: HOFFMANN, W. A. M; FURNIVAL, C. (Org.). **Olhar**: ciência, tecnologia e sociedade. São Carlos: Centro de CECHUFSCar, 2008.

RENDÓN ROJAS, M. A. **Bases teóricas y filosóficas de la Bibliotecología**. 2. ed. México-DF: UNAM, Centro Universitario de Investigaciones Bibliotecológicas, 2005.

RENDÓN ROJAS, M. A. La ciencia de la información en el contexto de las ciencias sociales y humanas. **DataGramZero**, v.9, n.4, ago. 2008.

SALAÛN, J. La redocumentarisation, un défi pour les sciences de l'information. **Études de Communication**, 30 [En ligne], mis en ligne le 01 octobre 2009. Disponível em <<http://edc.revues.org/index428.html>>. Acesso em 05 jun. 2014.

SMIT, J. W.; GUIMARÃES, J.A.C. Análise documentária. In: SMIT, J. W.; LIMA, Y. D. **Organização de arquivos**: XV curso de especialização. São Paulo: Escola de Comunicação e Artes/USP, 2001.

ANÁLISE DE DOMÍNIO HJØRLANDIANA SOB A LUZ DA FILOSOFIA DE DELEUZE

DOMAIN ANALYSIS HJØRLANDIAN UNDER THE LIGHT OF PHILOSOPHY OF DELEUZE

Igor Soares Amorim
Lígia Maria Arruda Café

Resumo: A Análise de Domínio é uma teoria fundamentada na relação direta entre comunidades discursivas e campos cognitivos, para sistematização do conhecimento e produção de sistemas de organização do conhecimento. Objetivo desse trabalho é interpretar a Análise de Domínio no contexto da pós-modernidade, optando pela Filosofia deleuziana como perspectiva. Pretende-se analisar, sob a óptica da Filosofia de Deleuze, o conceito de Análise de Domínio. A pesquisa é qualitativa, exploratória e bibliográfica. Utiliza-se a Análise de Conteúdo de Bardin como método de pesquisa. Para o *corpus* serão selecionados artigos de autoria de Hjørland encontrados bases *Library and Information Science Abstracts* (LISA) e *Web of Science e Scopus* (WoS) e no periódico *Advances in Knowledge Organization* (AKO). Adotando a técnica de leitura flutuante, estabeleceremos inferências e levantaremos índices para a aproximação. O mesmo procedimento será realizado em relação aos livros de Deleuze. Almeja-se, com a caracterização e o cruzamento dos índices da Análise de Domínio hjørlandiana e da Filosofia de Deleuze, pôr em diálogo as duas abordagens, de modo a endossar teoricamente a Ciência da Informação, assim como contextualizar a Organização do Conhecimento na pós-modernidade.

Palavras-chave: Análise de Domínio. Organização do Conhecimento. Gilles Deleuze. Pós-modernidade.

Abstract: The Domain Analysis is based on a direct relationship between discourse communities and cognitive fields for systematization of knowledge production and knowledge organization systems. This paper aims to interpret the Domain Analysis within the context of postmodernity, opting for Deleuzian's philosophy as perspective. It is intended to analyze, from the perspective of philosophy of Deleuze, the concept of Domain Analysis. The research is qualitative, exploratory and bibliographical. It uses Content Analysis of Bardin as a research method. For the corpus will be selected articles authored by Hjørland found bases Library and Information Science Abstracts (LISA) and Scopus and Web of Science (WoS) and in the journal *Advances in Knowledge Organization* (AKO). Adopting the technique of floating reading, will establish inferences and will raise indices for the approach. The same procedure will be performed in relation to Deleuze's books. One hopes, with the characterization and the intersection indices of hjørlandiana Domain Analysis and the Philosophy of Deleuze, bring into dialogue the two approaches, so theoretically endorse the Information Science, as well as contextualizing the Knowledge Organization in postmodernity.

Keywords: Domain Analysis. Knowledge Organization. Gilles Deleuze. Postmodernity.

1 INTRODUÇÃO

Mostafa (2013) traz um quadro com os trabalhos na Biblioteconomia e Ciência da Informação (BCI) que se utilizaram da Filosofia deleuziana, destacando as variadas possibilidades de frutificação que os trabalhos do filósofo francês podem render ao campo de

pesquisa. Seguindo essa linha de pensamento, acreditamos ser possível estabelecer conexões entre Deleuze e a BCI, por meio das ideias propostas por Hjørland¹⁸⁶. Tal reflexão será importante ao âmbito de pesquisa da área, não apenas como uma busca de abordagem inovadora, mas também pelo endossamento de fundamentos do campo. Nossa intervenção passa por repensar a Análise de Domínio (AD) pelo viés da Filosofia pós-moderna deleuziana e assim contribuir para as concepções de cunho social adotadas pela BCI.

Para Araújo (2003), a CI nasce muito próxima às Ciências Exatas, em seguida desenvolve um perfil crítico baseado no referencial teórico marxista (que passa a figurar em muitos trabalhos da área) e, posteriormente, sofre profundas alterações via os enfoques microsociológicos e interpretativos, que revelam uma tentativa da CI de se consolidar como pós-moderna. A AD é concebida nesta última tendência, por isso é uma hipótese coerente que ela apresente aspectos da Filosofia pós-moderna, assim como os de Ciência pós-moderna. Para Santos (1989), é um compromisso ético levar as Ciências ao status de pós-moderna, e é neste bojo que propomos relacionar a Filosofia de Deleuze e a Análise de Domínio hjørlandiana.

Diante do exposto, sintetizamos nossa questão de pesquisa: Até que ponto a AD apresenta-se em consonância com a Filosofia pós-moderna? A fim de respondê-la, propomos como objetivo geral: repensar o conceito de AD de Hjørland em uma perspectiva filosófica pós-moderna. E como objetivos específicos: a) Identificar as propriedades que sustentam o conceito de AD; b) Identificar na Filosofia de Deleuze princípios que possibilitem o diálogo crítico entre a filosofia pós-moderna e o conceito de AD; c) Interpretar o conceito de AD de Hjørland com base na Filosofia deleuziana.

2 LENDO UM DOMÍNIO COM HJØRLAND

Desde a década de 1990, Birger Hjørland é uma importante referência na área de BCI, devido a sua teoria da AD. Essa teoria se afasta da visão behaviorista e cognitivista predominantes na área até então, adotando um viés pragmático. Assim pauta-se na noção de comunidades discursivas, as quais produzem documentos e informações e se caracterizam por terem específicas necessidades informacionais. A AD visa não apenas fundamentar as práticas referentes as operações de classificação e representação da informação, mas também dar

¹⁸⁶ Quando o termo “Hjørland” não é acompanhado de uma especificação de ano, ele remete a integral produção teórico-conceitual desse autor à BCI, e não a um texto ou grupo de textos específicos.

consistência e maior rigor científico nos processos de representação do conhecimento e da informação (HJØRLAND, 2008).

O termo *Domain Analysis* foi concebido na engenharia de software e incorporado à BCI por Hjørland e Albrechtsen no texto de 1995, *Toward a new horizon in information science: domain-analysis*. A teoria posta à BCI, além de se apresentar como um método, aspira ser um paradigma (ou orientação de um), colocando em evidência os aspectos sociais que recobrem o fenômeno da informação.

Para Hjørland e Hartel (2003), a AD possui três dimensões: a Ontológica, a Epistemológica e a Sociológica – não podendo ser aplicada ou compreendida desconsiderando qualquer uma dessas. Para os autores, é por essas dimensões que se compreende a relação entre domínio geral e específico e seus desdobramentos no que tange aos princípios e estratégias de recuperação da informação. Para o dinamarquês, é a dimensão sociológica que viabiliza a empiria de sua teoria. Hjørland e Albrechtsen (1995) e Hjørland (2013) enfatizam o trabalho do sociólogo da ciência Richard Whitley, estadunidense que criou duas categorias para analisar o grau de desenvolvimento de uma disciplina, que dão conta tanto dos aspectos relativos ao conteúdo quanto aos estruturais (infraestruturais). As categorias são: institucionalização cognitiva e institucionalização social. Há que se destacar que, segundo Whitley existe uma estreita relação entre as duas facetas. É pautando-se em Whitley que Hjørland e Albrechtsen (1995) indicam três principais facetas para o levantamento de informações sobre um domínio, as três articulam-se entorno do conceito de comunidade discursiva: “estrutura de comunicação”; “padrões de comunicação” e “linguagem empregada no domínio”.

Hjørland e Albrechtsen (1995) afirmam tal perspectiva como inauguradora do paradigma social da CI, e ratificam o seu afastamento do positivismo e do racionalismo impregnados no paradigma cognitivista, em função de uma abordagem voltada mais ao funcionalismo e pragmatismo americano típicos do século XX (HJØRLAND; ALBRECHTSEN, 1995; HJØRLAND 2008; MOSTAFA, 2012).

Nesse sentido que para Wesolek (2012) as noções de “jogos de linguagem” de Wittgenstein e “domínio de conhecimento” são estreitamente relacionadas. Quando pensamos na identificação e representação de um domínio de conhecimento, consideramos que há jogos de linguagem nesse domínio que geram, por meio de negociações, um consenso. Portanto, a realização da análise de um domínio passa por identificar os jogos de linguagens e seus consensos. O consenso é fundamental na teoria hjørlandiana, pois é ele o conhecimento do domínio, o ente emergente do domínio. Por isso é tão central na AD a noção de comunidade

discursiva. É a comunidade quem discursa, que desenvolve o consenso, quem joga, quem constrói suas verdades demarcadas social e historicamente. A função da linguagem é essa, ela permitir que o social se mova nos discursos. Percebemos que, não só na AD, mas em toda a BCI, a linguagem é dotada de transversalidade.

3 DELEUZE E A AVENTURA DO PENSAMENTO

O filósofo francês Gilles Deleuze questiona-se sobre o que é, o que faz e o que produz, a Filosofia, a Ciência e a Arte. O que o filósofo fez ao longo de sua vida foi produzir pensamentos filosóficos. Em seus livros vemos isso, sempre criando novas perspectivas para abordar diversos temas, de uma maneira filosófica e inventiva. Para Deleuze, a Filosofia, a Ciência e a Arte compartilham uma questão, pois todas lidam com o pensamento. Pensar, na filosofia deleuziana, é se movimentar entre um virtual e um atual, o universo caótico e um mundo organizado, e este movimento as três áreas fazem, cada qual a seu modo. O pensamento só é criativo e inventor na Filosofia, na Ciência e na Arte, por isso não cabem nestas o ato de debater ou discutir, mas apenas criar! Essa criação aparece pela expressão do virtual que se dá num atual. (DELEUZE; GUATTARI, 2010)

A Ciência lida com o virtual. A ação científica dá-se pelo traçado de um plano de referência, o qual permite que a expressão científica apareça conferindo uma referência circunstancial num cruzamento de espaço-tempo. Assim, dizemos que a CI tem um plano (ou vários?). Toda ciência, seja exata, biológica, humana ou social, traça um plano de referência e cria funções. As funções lidam com corpos e o estados das coisas e são elas responsáveis pelo desaceleramento do caos por referências, atualizando os estratos científicos. O desaceleramento do caos gera resultantes, denominadas functivos, os quais dão densidade às funções. “A ciência é um sistema discursivo reflexivo e comunicativo, composto de funções e proposições, que lidam com o caos sempre o remetendo a referências.” (MOSTAFA; AMORIM; SOUZA, 2014) As teorias presentes na CI se formam assim, por meio de functivos e funções, que formam proposições e discursos de verdades num eixo paradigmático.

A Filosofia também vai ao caos, buscando expressar o seu infinito na sua dinamicidade. A Filosofia cria um plano do qual emanam componentes, por isso o plano é de imanência. Não cabe à Filosofia fazer a referência, demarcar um tempo-espaço, mas levantar componentes, percorrendo-os numa velocidade infinita, fazendo-os um conceito. Uma Filosofia da CI, então, necessariamente cria conceitos filosóficos sobre a informação. (DELEUZE; GUATTARI, 2010)

É importante sabermos que, para Deleuze, os planos de imanência, referência e de composição trocam intensidades, deslizando uns sobre os outros, estabelecendo relações entre as diferentes manifestações de pensamento. Portanto, é possível compreender a AD enquanto uma expressão científica, fundada em funções proposicionais, estabelecida sobre functivos e englobada em um plano de referência. Também pode-se fazer uma Filosofia da Análise do Domínio, identificando seus conceitos e planos em que se sustenta a teoria, como o conceito de tempo, de memória, de atualização, movimento, informação, etc. (DELEUZE; GUATTARI, 2010)

4 PROCEDIMENTOS METODOLÓGICOS

A pesquisa é qualitativa, bibliográfica e de caráter exploratória. Pretendemos realizar uma releitura da AD com base na Filosofia de Deleuze e, para tanto, usaremos a Análise de Conteúdo, proposta por Laurence Bardin (2003). Tal método compreende o tratamento e análise dos dados de forma sistemática, desde o levantamento e seleção de bibliografias e estruturação dos corpus de análise, até a etapa de tratamento de resultados e construção de inferências, sendo ele composto por três fases: a “pré-análise”, a “exploração do material” e o “tratamento dos resultados, inferências e interpretações”. Devido à recente aparição da AD na área, consideraremos não uma amostra, mas todos os artigos científicos recuperados nas bases de dados e na revista *Advances in Knowledge Organization* de autoria de Hjørland, portanto a pesquisa é censitária. Passaremos pelas três fase para alcançar os objetivos específico e, conseqüentemente, o objetivo geral deste trabalho.

O primeiro objetivo específico “identificar as propriedades que sustentam o conceito de AD” será alcançado por meio das seguintes etapas, que envolvem tanto a pré-análise como a exploração do material propostas por Bardin (2003):

Levantamento de artigos de periódicos em inglês de autoria de Birger Hjørland nas bases *Library and Information Science Abstracts* (LISA) e *Web of Science e Scopus* (WoS) e no periódico *Advances in Knowledge Organization* (AKO). Para esta etapa, buscaremos em cada volume do periódico textos do autor em questão. Na bases de dados utilizaremos como estratégia de busca o nome “Hjorland” no campo de autor. A busca será atemporal.

Por meio do método da leitura flutuante proposto por Bardin (2003), identificaremos, nos textos selecionados, fragmentos que registrem as propriedades do conceito de AD, categorizando-os em classes, possibilitando assim uma aproximação temática e a constituição dos índices de análise. Registraremos os dados no instrumento de coleta elaborados para esse fim.

O segundo objetivo “Identificar na Filosofia de Deleuze princípios que possibilitem o diálogo entre a filosofia pós-moderna e o conceito de AD” será alcançado por meio das seguintes etapas:

Levantaremos na Biblioteca Universitária da Universidade Federal de Santa Catarina livros do filósofo francês que possibilitem um diálogo com a proposta de Hjørland.

Por meio do método da leitura flutuante proposto por Bardin (2003), identificaremos, nos livros selecionados, extratos que registrem os princípios da Filosofia de Deleuze e que possam ser utilizados para refletir sobre o conceito de AD de Hjørland. Utilizaremos as classes e os dados coletados na primeira parte como orientação para identificação dos princípios em Deleuze, numa tentativa de aproximar os dois autores analisados na pesquisa. Registraremos os dados no instrumento de coleta elaborado para este fim.

O terceiro objetivo “Interpretar o conceito de AD de Hjørland com base na Filosofia deleuziana” será alcançado por meio do que Bardin (2003) denomina de tratamento dos resultados e inferências. Bardin (2003) afirma que, com o uso de sua proposta analítica, teremos um distanciamento necessário dos textos para, então, elaborarmos critérios desta análise. Com base nisso, proporemos uma discussão em torno dos pilares conceituais da AD hjørlandiana e dos princípios da Filosofia de Deleuze, entre os proeminentes autores neste início de século.

5 RESULTADOS ESPERADOS E CONSIDERAÇÕES PARCIAIS

Na LISA, subtraindo os artigos recuperados e não enquadrados no filtro pré-estabelecido, contabilizamos 44 artigos. Na WoS, foram 36. Os artigos comuns nas bases foram 24 artigos. Além desses, somam-se mais 3 artigos identificados na AKO. Desta forma, são 59 artigos referentes a produção de Hjørland a serem analisados.

Adotando a leitura flutuante (BARDIN, 2003), identificamos que “comunidade discursiva”, “domínio”, “conhecimento”, “linguagem”, “experiência e consenso”, “informação e informatividade” são noções chaves para a AD. A fim de dialogar com princípios filosóficos deleuzianos, identificamos os seguintes termos no trabalho do francês pós-moderno: “rizoma”, “agenciamento”, “devir”, “acontecimento”, “Corpo sem Órgãos”. A partir desses termos estabeleceremos um diálogo entre os dois discursos, renovando a teoria de Hjørland na Organização do Conhecimento, conferindo ao campo aspectos filosóficos e científicos da pós-modernidade. O diálogo se dará por meio do cruzamento entre todos os índices levantados referentes à AD com todos os índices selecionados dos livros de Deleuze.

Almejamos contribuir para a “atualização” da Organização do Conhecimento com fundamentação teórica frente as novas características da sociedade pós-moderna; tomando como base a caracterização e o diálogo da AD de Hjørland com elementos da Filosofia de Deleuze. Dessa forma lançaremos uma outra perspectiva na CI, com uma outra possível fundamentação, de cunho deleuziana, pela da teoria da AD. Para tanto, pretendemos realizar deslizamentos entre os planos de imanência e de referência, com a intenção de renovar a CI e quiçá criar conceitos filosóficos ou funções ao campo.

REFERÊNCIAS

- ARAÚJO, C. A. A. Ciência da Informação como Ciência Social. **Ciência da Informação**, v. 32, n. 3, p. 21-27, 2003.
- BARDIN, L. **Análise de Conteúdo**. Lisboa. Edições 70, 2003.
- DELEUZE, G.; GUATTARI, F. **O que é Filosofia?**. 3. ed. Rio de Janeiro: Editora 34, 2010.
- HJØRLAND, B. Theories of Knowledge Organization: theories of Knowledge. **Knowledge Organization**, v. 40, n. 3, p. 169-181, 2013.
- HJØRLAND, B. What is Knowledge Organization. **Knowledge Organization**, v. 35, n. 2/3, p. 86-101, 2008.
- HJØRLAND, B.; HARTEL, J. Afterword: Ontological, epistemological and sociological dimensions of domains. **Knowledge Organization**, v. 30, n. 3/4, p. 239-245, 2003.
- HJORLAND, B; ALBRECHTSEN, H. Toward a new horizon in Information Science: Domain-Analysis. **Journal of the American Society for Information Science**, v. 46, n. 6, p. 400-425, 1995.
- MOSTAFA, S. P. Charles Pierce, Gilles Deleuze e a Ciência da Informação. **Informação & Sociedade**, v. 22, n. 1, p. 27-37, 2012.
- MOSTAFA, S. P. **Filosofia da diferença e a ciência da informação**. Rio de Janeiro: E-papers, 2013.
- MOSTAFA, S. P.; AMORIM, I. S.; SOUZA, L. M. A. Filosofia e discurso na Ciência da Informação: tessitura de encontros. **Filosofia da Informação**, 2014. [No prelo].
- SANTOS, B. S. **Introdução a uma ciência pós-moderna**. Rio de Janeiro: Graal, 1989.
- WESOLEK, A. Wittgensteinian support for domain analysis in classification. **Library Philosophy and Practice**, p. 1-10, 2012.

METODOLOGIAS, INSTRUMENTOS E INTERFACES DE ORGANIZAÇÃO DA IMAGEM NA WEB NO CONTEXTO DO SISTEMA “MEMÓRIA VIRTUAL”

METHODS, INSTRUMENTS AND INTERFACE OF THE ORGANIZATION OF IMAGE ON THE WEB IN THE CONTEXT OF SYSTEM “VIRTUAL MEMORY”

Letícia Reis da Silveira
Luciana de Souza Gracioso

Resumo: O estudo identificou em que medida alguns dos sistemas online disponíveis para organização e recuperação de fotografias, contemplam em sua interface de busca para o usuário, algumas das principais categorias relacionadas ao tratamento temático de fotografias, tal como pesquisados e indicados na literatura da Ciência da informação brasileira, mas que foram desenvolvidas a partir de referencial metodológico internacional. Dentre estes sistemas, especial atenção foi dada à análise do Memória Virtual, software desenvolvido no bojo do Projeto “Critérios e Metodologias para a realização de inventário do Patrimônio Cultural Rural Paulista”, da Universidade Federal de São Carlos, Brasil, que propõe a construção de um padrão de descrição da informação contemplando campos específicos para os processos de indexação de bens patrimoniais. Para coleta de dados, foi desenvolvido e aplicado roteiro de avaliação sobre tratamento temático de imagens em sistemas voltados ao patrimônio e memória. A partir da aplicação do roteiro foi possível concluir que este já é um sistema muito completo e explora de forma ampla todos os níveis necessários para descrever minuciosamente o bem patrimonial, inclusive quando se trata de fotografias.

Palavras-chave: Organização do conhecimento. Imagem. Memória.

Abstract: The study identifies to what degree some of the online systems available for organizing and retrieving photos include some of the main categories related to thematic processing of photographs in their user search interface, as researched and reported in the literature of Brazilian Information Science, but which were developed from an international methodological framework. Among these systems, special attention was given to the analysis of Virtual Memory, software that has been developed within the project "Criteria and Methodologies to perform the Paulista Cultural Rural Heritage Inventory", from Federal University of São Carlos, Brazil, which proposes the construction of a standard description of information covering specific fields for the processes of indexing cultural heritage images. For data collection, it was developed and implemented a roadmap for the evaluation of thematic treatment of images in systems focused on the heritage and memory. The implementation of the guideline for analyzing the Virtual Memory software enabled to conclude that this is a very complete system, which broadly explores all the levels needed to thoroughly describe cultural heritage, including photographs, positively responding to the guideline prepared for this work, with some exceptions, such as the Expressive Dimension field and Social Indexing.

Keywords: Knowledge Organization. Image. Memory.

1 INTRODUÇÃO

O estudo identificou em que medida alguns dos sistemas online disponíveis para organização e recuperação de fotografias, contemplam em sua interface de busca para o usuário, algumas das principais categorias relacionadas ao tratamento temático de fotografias,

tal como pesquisados e indicados na literatura da Ciência da informação brasileira, mas que foram desenvolvidas a partir de referencial metodológico internacional. Dentre estes sistemas, especial atenção foi dada à análise do Memória Virtual, software desenvolvido no bojo do Projeto “Critérios e Metodologias para a realização de inventário do Patrimônio Cultural Rural Paulista”, da Universidade Federal de São Carlos, Brasil, que propõe a construção de um padrão de descrição da informação contemplando campos específicos para os processos de indexação de bens patrimoniais. Para coleta de dados, foi desenvolvido e aplicado roteiro de avaliação sobre tratamento temático de imagens em sistema voltado ao patrimônio e memória. A partir da aplicação do roteiro para foi possível concluir que este já é um sistema muito completo e explora de forma ampla todos os níveis necessários para descrever minuciosamente o bem patrimonial, inclusive quando se trata de fotografias.

Configuraram como objetivos da pesquisa: identificar e analisar na literatura científica de Ciência da Informação, metodologias relacionadas ao tratamento temático de fotografias e a partir disto, desenvolver um roteiro de observação e verificação junto a sistemas de informação disponíveis online voltados à organização e disponibilização de fotografias; selecionar, a partir de critérios *Globais, Locais e Institucionais*, sistemas de informação online direcionados a organização e disseminação de fotografias, sendo eles: Corbis, Fundação Casa de Rui Barbosa, *Visual Information Access*, Software de Apoio à Comunicação Integrada (SACI), Dedalus, Fundo Florestan Fernandes e Flickr; aplicar o roteiro desenvolvido em tais sistemas, tanto para analisá-los como para identificar novas categorias de tratamento temático de fotografias que têm sido utilizados na *web* e que não foram identificados na literatura científica da Ciência da Informação brasileira selecionada para elaboração do roteiro. A partir da aplicação do roteiro nas interfaces *web* e com a identificação de novas categorias utilizadas para tratamento da informação nos sistemas analisados, este instrumento foi complementado e a partir disto, aplicado para a análise do Memória Virtual, sendo esta a interface sobre a qual foi focada maior atenção da pesquisa, uma vez que se pretende poder contribuir diretamente com o desenvolvimento deste sistema sinalizando possíveis intervenções ou fortalecendo as orientações já desenvolvidas e aplicadas no trato da fotografia.

Metodologicamente, o trabalho se configurou enquanto pesquisa exploratória bibliográfica e documental, desenvolvendo em parte ações de pesquisa-intervenção e se utilizando de observação participante para coleta de dados. Tendo em vista o movimento de retroalimentação do roteiro de observação que foi desenvolvido para análise dos sistemas, optou-se por já apresentar, na sequência desta introdução, o detalhamento do percurso metodológico desenvolvido.

2 MÉTODOS DE PESQUISA

O presente trabalho teve como objetivo geral estudar metodologias de organização de imagens na área da Ciência da Informação e verificar se os sistemas de informação atuais disponíveis na Web, que lidam com a informação imagética, põem em prática as orientações advindas da literatura científica da área no que diz respeito ao tratamento temático de fotografias. Especificamente, teve-se como objetivo criar, a partir destas metodologias, um instrumento para poder analisar alguns destes sistemas, o qual foi chamado “roteiro de observação e análise”. Além da elaboração do instrumento de pesquisa, o trabalho pretendeu fazer a caracterização dos sistemas de informação selecionados, aplicando o roteiro nas interfaces de busca e recuperação de dados do usuário, realizando uma análise a partir da perspectiva do usuário, e obtendo novos parâmetros que poderiam não ter sido sugeridos nas metodologias estudadas. Ainda como objetivo específico, o trabalho visou aplicar o roteiro ao software Memória Virtual como tentativa de validar as categorias elaboradas para a descrição de bens patrimoniais (no entanto, com foco nas fotografias), verificando se elas atendem às principais práticas sugeridas pela literatura e, caso necessário, oferecer parâmetros que possam auxiliar no tratamento e recuperação das imagens pelo sistema.

Quanto à modalidade de pesquisa, o estudo foi definido como exploratório, pois pretendeu, ao entrar no universo dos estudos de imagens, analisar e tentar definir as melhores práticas quanto o tratamento deste tipo informacional para a disponibilização aos usuários e, conjuntamente, ao identificar tais práticas nos sistemas selecionados. Além de exploratória, a pesquisa também pode ser definida como uma pesquisa-intervenção, pois como dito, tem como um de seus objetivos a análise e possível sugestão de parâmetros para aprimoramentos no Memória Virtual. É ainda considerada uma pesquisa bibliográfica, pois visou em um primeiro momento, recuperar em bases de dados nacionais e internacionais o conhecimento científico sobre um determinado tema, ou seja, a organização da informação com foco em imagens.

Como em um segundo momento foi realizada observação participante do autor junto às interfaces do usuário dos sistemas *web*, configurando o eixo principal da análise, e tendo por pressuposto que essas interfaces podem ser consideradas documentos, ela também pode ser considerada como uma pesquisa documental. É importante ressaltar que este estudo teve um corte transversal, uma vez que a seleção dos sistemas se deu em apenas um momento da pesquisa.

É necessário explicitar que quanto à aplicação do roteiro de observação, foi feito um recorte em relação às análises. Foram analisados três eixos principais: o recorte acadêmico

(sistemas desenvolvidos por instituições de ensino superior), o recorte institucional (sistemas desenvolvidos por instituições que lidam com informação patrimonial), e o recorte social (sistemas desenvolvidos para cadastro de fotografias no ambiente *web*). Todas essas instituições foram analisadas em âmbito global, nacional e local.

3 RESULTADOS

O movimento de aplicação do roteiro para a análise dos sistemas acabou por proporcionar também uma perspectiva a respeito da Indexação Social no sentido de observar como ela se apresenta cada vez mais forte como iniciativa de organização da informação imagética em sistemas conceituados em nível nacional e internacional. Por este motivo, julgou-se necessário complementar o instrumento de pesquisa elaborado, acrescentando a ele a categoria a respeito da abertura dos sistemas à Indexação Social, situando-a no segmento “Representação da Informação”,

O QUADRO 1 apresenta uma síntese dos resultados da aplicação do roteiro aos sistemas escolhidos para a análise já incluindo os resultados da aplicação ao Memória Virtual. Como é possível observar, as categorias estão separadas entre as três dimensões de análise e as siglas “S” e “N” significam “Sim” e “Não”, respectivamente:

QUADRO 1 - Quadro-síntese contendo os resultados da aplicação do roteiro aos sistemas analisados

DIMENSÃO	CATEGORIAS	Corbis	Fund. Casa Rui Barbosa	SACI	VIA	De alius	Fundo Florestan Fernandes	Flickr	Memória Virtual
		ANÁLISE DE CONTEÚDO	Permite cadastro livre de fotografias?	N	N	N	N	N	N
A imagem está contextualizada?	S		S	S	S	N	S	S	S
Existe campo para a descrição informativa da imagem?	S		S	S	S	S	S	S	S
Existe campo para descrição da dimensão expressiva da imagem?	N		S	N	N	N	N	S	N
Existe campo para resumo do conteúdo da imagem?	S		S	N	N	S	S	S	S
Existe campo para a contextualização história da imagem?	N		N	N	S	S	S	S	S
Existem informações sobre condições de acesso ao material?	N		S	S	S	S	S	S	S
REPRESENTAÇÃO DA INFORMAÇÃO	As imagens são indexadas por assuntos e/ou palavras-chaves?	S	S	S	S	S	S	S	S
	Existe campo para a indexação social?	S	S	N	N	N	S	S	N
	Existe campo para a legenda da imagem?	S	S	S	S	S	S	S	S
	Há uso de linguagem controlada para a indexação?	N	S	N	S	S	N	N	S
RECUPERAÇÃO DA INFORMAÇÃO	Existe ferramenta de busca?	S	S	N	S	S	S	S	S
	O usuário pode realizar busca avançada/integrada?	S	S	N	S	S	S	S	N
	A imagem é recuperada com boa resolução?	S	S	S	S	N	S	S	S

Fonte: os autores

4 CONSIDERAÇÕES FINAIS

Os objetivos estabelecidos para este trabalho foram os de reunir e caracterizar as metodologias da Ciência da Informação publicadas e citadas em maior frequência no contexto das publicações brasileiras, a respeito de análise e organização de fotografias e, a partir de orientações advindas destas metodologias, elaborar um roteiro de observação e análise de sistemas de organização de fotografias disponíveis na Web, no intuito de verificar em que medida os sistemas vigentes atendem os pré-requisitos de tratamento temático da imagem preconizados nas metodologias estudadas. O roteiro elaborado foi aplicado em sete sistemas *web* que tratavam do conteúdo imagético para verificar quais categorias vem sendo desenvolvidas e utilizadas por esses sistemas em relação à análise e organização de imagens, e também verificar se estes sistemas atuais contemplam as técnicas sugeridas pela literatura científica da área da Ciência da Informação no que diz respeito à organização e representação da informação e também aspectos quanto às suas ferramentas de busca.

A partir da aplicação do roteiro para a análise do software Memória Virtual, foi possível concluir que este já é um sistema muito completo e explora de forma ampla todos os níveis necessários para descrever minuciosamente o bem patrimonial, inclusive quando se trata de fotografias, respondendo de forma praticamente totalmente positiva ao roteiro elaborado por este trabalho, com algumas exceções, como no caso do campo para Dimensão Expressiva e espaço para Indexação Social.

Quanto à Indexação Social, este trabalho defende que ela seja levada em conta, pois a perspectiva sócio cognitiva do usuário no momento de indexar ou buscar por suas necessidades informacionais, de fato, podem contribuir, e muito, para uma recuperação mais eficiente dos resultados das buscas. Uma sugestão seria utilizar essa perspectiva dos usuários de forma indireta, sendo que eles poderiam ajudar na indexação de termos a partir dos resultados de suas buscas. Caso o sistema seja capaz de armazenar os termos de busca utilizados pelos usuários, tais termos poderiam ser recuperados e posteriormente analisados por um profissional catalogador, ou revisor, e adicionados aos campos de assuntos e descritores dos bens patrimoniais aos quais estiverem relacionados.

Como todos os estudos, o atual trabalho também acabou deparando-se com algumas limitações e desafios metodológicos, como no caso da dificuldade em construir instrumentos para a pesquisa e posteriormente delimitar as metodologias utilizadas, principalmente quanto a chamar este estudo de pesquisa-intervenção. Outra dificuldade seria em relação ao número de sistemas e fotografias analisados. Nos dias atuais, muitos são os sistemas desenvolvidos para o trato de fotografias, e milhares são as fotografias indexadas por eles. No entanto, o

tempo relativamente curto disponível para a realização da pesquisa não permitiu que se pudesse fazer uma análise mais aprofundada, com mais sistemas a analisar, e mais fotografias, o que acaba por abrir portas para a continuidade desta pesquisa ou o desenvolvimento de novas relacionadas ao tema, aplicando o roteiro a mais sistemas, a diferentes contextos e até mesmo testando sua eficácia, por exemplo.

Para finalizar, conforme o que foi exposto, a informação, seja ela em forma de texto, imagem, áudio ou som existe em forma abundante e necessita ser transmitida para que faça sentido e possa ajudar o usuário com a formação de conhecimento para tomada de decisões, preencher lacunas informacionais e até mesmo para o resgate e preservação de sua Memória. A informação imagética, por sua vez, cada vez mais ganha espaço como fonte de informação, transmitindo conhecimento como um todo ou de forma complementar a outra fonte de informação. Como demonstrado neste trabalho, a preocupação cada vez maior por parte de instituições em formar acervos e disponibilizá-los ao público comprova este fato. Desta forma, é de extrema importância levar em consideração as melhores práticas para o tratamento desta informação para que possa chegar de forma completa e satisfatória ao usuário.

Com os esforços realizados neste trabalho e a construção do roteiro de observação e análise como uma ferramenta de diagnóstico, vislumbra-se que este possa ser útil para outras instituições e pesquisas no que diz respeito a processos de tomada de decisões e auxílio em práticas pertinentes para a análise e organização de imagens, principalmente no que diz respeito às de caráter histórico.

REFERÊNCIAS

ABREU, J. G. de; MONTEIRO, S. D. Matrizes da linguagem e a organização virtual do conhecimento. **Ciência da informação**, v. 139, n. 1, 2009.

COSTA, L. S. F.; SCARPELINE, R.; NAKAGAWA, E. Y. Uma proposta teórico-metodológica de inventário patrimonial no contexto das fazendas históricas paulistas. In: Seminário: Preservação do patrimônio agroindustrial, 2º - Lugares de Memória. **Anais...** São Carlos, Escola de Engenharia de São Carlos – USP, 2010. CD-ROM.

KOSSOY, B. Construção e desmontagem da informação fotográfica: teoria e história. **Revista Acadêmica Ciências Agrárias e Ambientais**, São Paulo, p. 224-232, 2004.

MAIMONE, G. D. **Estudo do tratamento informacional de imagens artístico-pictóricas: cenário paulista – análises e propostas**. Dissertação (mestrado). Pontifícia Universidade Católica de Campinas. Campinas, 2007. 140 f.

MANINI, M. P. **Análise documentária de fotografias: um referencial de leitura de imagens fotográficas para fins documentários**. Tese (doutorado). Universidade de São Paulo. São Paulo, 2002. 231 f.

MARTINEZ, L. V. **NOBRADE e a descrição de fotografias nas diversas áreas da Ciência da Informação**. Trabalho de Conclusão de Curso (graduação). Universidade Federal do Rio Grande do Sul. Porto Alegre, 2009. 76 p.

TOREZAN, I. M. V. **Fotografia e Informação: aspectos gerais de análise e indexação da imagem**. Dissertação (mestrado). Universidade de Brasília. Brasília, 2007. 121 f.

ORGANIZAÇÃO DA INFORMAÇÃO E DO CONHECIMENTO DE DOCUMENTOS ARTÍSTICOS: INTEGRANDO MUSEU, BIBLIOTECA E ARQUIVO

INFORMATION AND KNOWLEDGE ORGANIZATION OF ARTISTIC DOCUMENTS: INTEGRATING MUSEUM, LIBRARY AND ARCHIVE

Giovana Deliberali Maimone

Resumo: A existência de correlações entre algumas das atividades efetuadas no ambiente museológico, que em muitos casos consiste de biblioteca, arquivo e o próprio museu, possibilita pensar a integração funcional destes setores já que, embora contextualizados de modo diferente quanto ao uso e ambiência destes espaços, o propósito é o mesmo - tratar a informação de modo que esteja disponível aos gestores, funcionários e usuários da informação museológica. Neste sentido, tem-se como objetivo primordial apresentar, debater e propor aperfeiçoamentos às ferramentas utilizadas pela Ciência da Informação que proporcionam otimização dos processos documentários ocorridos no ambiente museológico. A metodologia desta pesquisa consiste em trabalhos teóricos e práticos, estando subdivididas em cinco etapas: Pesquisa Bibliográfica, Análise de Campo, Desenvolvimento de Propostas, Aplicação destas Propostas e Elaboração do Relatório Final da Pesquisa, que será realizada no período de dois anos. Espera-se que os resultados ofereçam contribuições significativas para a Organização da Informação e do Conhecimento no que diz respeito aos documentos artísticos, tratados no museu, no arquivo e na biblioteca.

Palavras-chave: Ambiente Museológico. Documentos Artísticos. Organização da Informação e do Conhecimento. Processos Documentários.

Abstract: The existence of correlations between some of the activities done in the museum environment, which in most cases consists of library, archive and museum itself thinking enables the functional integration of these sectors as although contextualized differently about the use and ambiance of these spaces, the purpose is the same - so treat the information that is available to managers, staff and users of museum information. In this sense, it has the primary objective to present, discuss and propose improvements to the tools used by information science documentaries that provide optimization of processes occurring in the museum environment. The methodology of this research consists of theoretical and practical work, and is subdivided into five stages: Bibliographical Research, Field Analysis, Proposal Development, Implementation of these proposals and Preparation of the Final Research Report, which will be held in two years. It is expected that the results provide significant contributions to the Organization of Information and Knowledge with regard to artistic documents processed in the museum, in archive and library.

Keywords: Museum Environment. Artistic Documents. Information and Knowledge Organization. Documentaries Processes.

1 INTRODUÇÃO

A organização da informação e do conhecimento, considerada uma das subáreas da Ciência da Informação, está voltada ao estudo de processos que possibilitem o acesso aos registros do conhecimento (documentos diversos –, livros, pinturas, filmes, *e-books* etc.), a fim de que se tenha uma recuperação informacional adequada dos itens buscados.

Neste sentido, a existência de correlações, ainda não suficientemente identificadas e harmonizadas, entre algumas destas atividades no ambiente do acervo museológico, bibliográfico e arquivístico sinalizam a possibilidade efetiva de uma sinergia na organização e disseminação da informação. Dito de outro modo, historicamente, a diferença marcou a identidade dos objetivos e do trabalho desses equipamentos culturais que redundaram em parcializações dos sistemas de informação. Na contemporaneidade, com a ideia de fluxo subsumindo os processos de troca, não só duplicam-se os tratamentos mas também distanciam-se, às vezes de modo incontornáveis, conteúdos que mantêm forte vínculo associativo, independente do suporte, do lugar em que se encontram ou da coleção que integram.

Assim, a integração museu / biblioteca / arquivo no interior dos equipamentos museológicos é de extrema importância, uma vez que permite a obtenção de resultados mais completos para a curadoria do museu na organização de exposições, para a biblioteca como fonte de referência e informação, para o público que o visita, pela percepção de complementaridade das obras e, para a gestão integral do museu.

A obra material e suas diferentes formas de representação se relacionam, possibilitando a observação da associação entre acervo museológico, bibliográfico e arquivístico – sob um ponto de vista dinâmico e mais eficiente em nível comunicacional. Segundo Smit (2010, p. 34)

um museu não é composto unicamente do acervo museológico, pois ele ainda necessita de documentos de gestão do acervo, de apoio para a pesquisa e para seu funcionamento institucional.

Dentre os diferentes tipos de documentos (aqui denominados genericamente de documentos artísticos) presentes nos museus, há interesse, em particular: pelo conjunto do acervo museológico, que contém os objetos (peças, pinturas, esculturas); do acervo arquivístico e do acervo bibliográfico, que são compostos de materiais que servem para apoiar pesquisas desenvolvidas no museu. Esse último justifica-se por sua função de complemento das necessidades informacionais de artistas, pesquisadores, curadores, do público genérico interessado, entre outros, funcionando como fonte de informações (SMIT, 2010).

Independentemente do uso (ou não) que se faça da informação, esta pesquisa preocupa-se em tentar “democratizar” seu acesso, a fim de que satisfaça, de forma crescente, as necessidades do maior número possível de usuários (internos e externos¹⁸⁷), garantindo-se a

¹⁸⁷ Por usuário interno entende-se indivíduos que trabalham no museu independentemente do setor. Já os usuários externos constitui-se do público visitante / utilizador destes espaços.

geração contínua de conhecimentos. Entende-se, ainda que a democratização do acesso não se faz apenas através de facilitadores que podem apenas nivelá-lo, mas sim através da diversidade de pontes que permitam o trânsito da complexidade de todo artefato humano.

Inserido nesta temática, o campo da Ciência da Informação pode contribuir significativamente investigando instrumentos para aperfeiçoamento comunicativo como são as Linguagens Documentárias e ainda analisando tipos diferentes de “documentação” realizadas sobre as diversas modalidades de materiais encontrados nos museus (pinturas, objetos, livros, catálogos, ficha de procedência, custo da obra, datas de restauração, etc.). Visando a discussão de ideias dos autores de áreas correlatas, porém diversas, como a Museologia, a Arquivologia e a Ciência da Informação, e a fim de aprofundar reflexões e aplicações sobre teorias e métodos propomos o desenvolvimento do trabalho intitulado ORGANIZAÇÃO DA INFORMAÇÃO E DO CONHECIMENTO DE DOCUMENTOS ARTÍSTICOS: INTEGRANDO MUSEU, BIBLIOTECA E ARQUIVO.

A pesquisa proposta teve origem em duas questões centrais:

os acervos mostrados e organizados (tratados informacionalmente) pelos museus funcionam efetivamente como fontes de acesso ao conhecimento, ou seja, os métodos são adequados para comunicar informação?

a Ciência da Informação não poderia auxiliar de alguma maneira (com instrumentos organizativos, resumos, índices, estudos narrativos ou outros) nessas atividades, substanciais para o encontro usuário – informação?

Como objetivo geral propõe-se apresentar as ferramentas utilizadas pela Ciência da Informação para tratamento do material artístico (tesauros, listas de classificação, fontes terminológicas, etc.) e propor aperfeiçoamentos que proporcionem otimização dos processos documentários ocorridos no ambiente museológico.

E como objetivos específicos, temos:

- Identificar a relação entre a descrição de obras de arte (objetos, pinturas, instalações, etc.), de documentos arquivísticos e a de materiais impressos (livros, catálogos, periódicos, etc.).
- Contextualizar os diferentes usos e ambiências dos documentos artísticos procurando delimitar fronteiras de aplicação documentária e encontrar um elo comum de relação entre acervo museológico, arquivístico e bibliográfico.
- Definir os termos: público, usuários e visitantes de museus para direcionar o enfoque das atividades de tratamento documentário.

- Comparar linguagens documentárias e ferramentas afins já existentes na área de artes e propor sugestões de aprimoramento, quando possível.

2 ORGANIZAÇÃO DA INFORMAÇÃO E DO CONHECIMENTO DE DOCUMENTOS ARTÍSTICOS

Os estudos e aplicações em Organização da Informação e do Conhecimento possibilitam aperfeiçoar metodologias e instrumentos comunicacionais de diversas áreas a partir do estudo linguístico e terminológico que estão incluídos na chamada “Linguística Documentária”.

Deriva-se dessa conjectura as seguintes hipóteses:

- a descrição de objetos (museológicos, arquivísticos e bibliográficos) pauta-se nos objetivos das instituições, nos tipos de materiais disponíveis e ainda no interesse, mais ou menos específico dos utilizadores desses acervos, fato que exige tratamento documentário diferenciado, porém complementar.
- é preciso dialogar para atingir a função social dos equipamentos culturais. Então, uma vez que a diversidade existe, trabalhamos com os conteúdos para buscar a complementaridade, a sinergia informacional e os benefícios que as atividades documentárias podem oferecer aos mais diversos tipos de materiais.
- assim como ocorre na biblioteca e no arquivo, no museu os usuários estão interessados em saciar suas necessidades de informação para aperfeiçoar seu conhecimento, e para isso, necessitam de informações normalizadas e específicas.
- os museus, arquivos e bibliotecas de ambientes museológicos podem utilizar das investigações da Ciência da Informação (Organização da Informação e do Conhecimento / Linguística Documentária / Linguagens Documentárias) para solucionar problemas de informação ou colaborar com serviços informacionais mais funcionais e especializados comunicativamente.

Para avaliação das hipóteses acima apresentadas e valendo-se do acesso disponibilizado pela Universidade de São Paulo à artigos e bases de dados nacionais e internacionais será investigado, no âmbito da Ciência da Informação, os modos de tratamento documentário que poderiam ser aplicados à área de artes, tendo como referências as linguagens documentárias e correntes teóricas e aplicadas correlatas.

O desenvolvimento do trabalho se dará em três etapas principais, sendo:

Primeira etapa: Pesquisa Bibliográfica - levantamento e análise de referencial teórico da temática investigada, utilizando-se como fontes de pesquisas:

- Livros;
- Artigos de periódicos;
- Bases de dados: LISA (*Library and Information Science Abstracts*), ISTA (*Information Science & Technology Abstracts*), BRAPCI (Base de Dados Referencial de Artigos de Periódicos em Ciência da Informação), SciELO;
- Biblioteca Digital de Teses e Dissertações;
- *Art & Architecture Thesaurus*;
- *Thesauri & Controlled Vocabularies - Library of Congress*;
- Sistemas de Classificação Bibliográfica: CDD e CDU;
- Norma ANSI/NISO Z39.19-2005 (*Guidelines for the Construction, format, and Management of Monolingual Controlled Vocabularies*);
- ISO 25964 – *The International Standard for Thesauri and Interoperability with other Vocabularies*;
- Normas do CIDOC¹⁸⁸-ICOM¹⁸⁹

Observação: Esta etapa será constante durante todo o processo de pesquisa, pois há necessidade de recorrer à teoria para fortalecer o pensamento em formação / transformação.

Segunda etapa: Desenvolvimento de propostas metodológicas de aperfeiçoamento documentário em ambiente museológico.

Terceira etapa: Elaboração de relatório final de pesquisa pautada em argumentos de caráter conclusivo a partir das análises teóricas e práticas realizadas.

Técnicas de pesquisa:

- Análise de conteúdo do *corpus*;
- Comparação de instrumentos e ferramentas, demonstrando aspectos positivos e negativos;
- Propostas de aperfeiçoamento que integrem acervos museológicos, arquivísticos e bibliográficos;
- Desenho dos resultados conseguidos.

¹⁸⁸ CIDOC – International Committee for Documentation.

¹⁸⁹ ICOM – International Council of Museums.

3 CONSIDERAÇÕES GERAIS

Espera-se que os resultados ofereçam contribuições significativas para a Organização da Informação e do Conhecimento no que diz respeito aos documentos artísticos, tratados no museu, no arquivo e na biblioteca. No campo do desenvolvimento científico está previsto: analisar e desenvolver proposta(s) metodológica(s) de aperfeiçoamento documentário em ambientes museológicos, sempre recorrendo à Ciência da Informação. Do ponto de vista prático: publicar artigos de periódicos, trabalhos em eventos científicos (nacionais e internacionais) da área e consolidar proposta para o desenvolvimento de uma série de publicações sobre "Metodologias de Organização da Informação e do Conhecimento de Documentos Artísticos".

REFERÊNCIAS

- AAT. Art & Architecture Thesaurus. Los Angeles, EUA: The Getty Research Institute, 2013. Disponível em: <<http://www.getty.edu/research/tools/vocabularies/aat/>>.
- ANSI/NISO **Z39.19**: guidelines for the construction, format, and management of monolingual controlled vocabularies. Bethesda: NISO Press, 2005. 172 p. Disponível em: <http://www.niso.org/apps/group_public/download.php/6487/Guidelines%20for%20the%20Construction,%20Format,%20and%20Management%20of%20Monolingual%20Controlled%20Vocabularies.pdf> .
- AUSTIN, D.; DALE, P. Diretrizes para o estabelecimento e desenvolvimento de tesouros monolíngües. Brasília: IBICT/SENAI, 1993.
- BRUNO, M. C. O. (Coord.). **O ICOM-Brasil e o pensamento museológico brasileiro: documentos selecionados**. São Paulo: Pinacoteca do Estado: Secretaria de Estado da Cultura: Comitê Brasileiro do Conselho Internacional de Museus, 2010. v. 1, p. 239.
- BRUNO, M. C. O. (Coord.). **O ICOM-Brasil e o pensamento museológico brasileiro: documentos selecionados**. São Paulo: Pinacoteca do Estado: Secretaria de Estado da Cultura: Comitê Brasileiro do Conselho Internacional de Museus, 2010. v. 2, p. 164.
- CAMARGO-MORO, F. de. **Museu: aquisição/documentação**. Rio de Janeiro: Livraria Eça Editora, 1986.
- CAPURRO, R.; HJØRLAND, B. O conceito de informação. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 12, n. 1, p. 148 – 207, jan. / abr. 2007.
- CIDOC. **Home**. Disponível em: <<http://network.icom.museum/cidoc/>>. Acesso em: 28 jul. 2013.
- CINTRA, A. M. M. *et al.* **Para entender as linguagens documentárias**. 2. ed. rev. e ampl. São Paulo: Polis, 2002. 92 p. (Coleção Palavra-Chave, 4).
- DAHLBERG, I. Teoria do conceito. **Ciência da Informação**, Rio de Janeiro, v. 7, n. 2, 1978. p. 101 – 107.
- DUBOIS, J. *et al.* **Dicionário de linguística**. São Paulo: Cultrix, 1973. 653 p.

FERREZ, H. D. Documentação museológica: teoria para uma boa prática. In: CADERNOS de Ensaio, n. 2, Estudos de Museologia. Rio de Janeiro: MinC/IPHAN, 1994.

GARDIN, J-C. Document analysis and linguistic theory. **Journal of documentation**, v. 29, n. 2, jun. 1973.

HARPRING, P. **Introduction to controlled vocabularies**: terminology for Art, Architecture, and Other Cultural Works. Los Angeles: Getty Research Institute, 2010. 244 p.

HJØRLAND, B. What is Knowledge Organization (KO)? **Knowledge Organization**, v. 35, n. 2/3, p. 86 – 101, 2008.

ICOM (Brasil). **Home**. Disponível em: <<http://icom.org.br/>>. Acesso em: 28 jul. 2013.

ICOM. **Home**. Disponível em: <<http://icom.museum/>>. Acesso em: 28 jul. 2013.

ISO 25964: The International Standard for Thesauri and Interoperability with other Vocabularies (2011, 2013).

ISO 2788/BS5723 (1986/87). **Guidelines to establishment and development of monolingual thesauri**. Genève; London: International Standard Organization; British Standard.

LE COADIC, Y.-F. **A Ciência da Informação**. 2. ed. rev. e atual. Brasília, DF: Briquet de Lemos/Livros, 2004. 124 p.

OTLET, P. **Traité de documentation**: le livre sur le livre. Bruxelles: Editions Mundaneum, 1934. . Disponível em: <http://lib.ugent.be/fulltxt/handle/1854/5612/Traite_de_documentation_ocr.pdf>. Acesso em: 02 fev. 2010.

SAGER, J. C. La terminología: puente entre varios mundos. In: CABRÉ, M. T. **La terminología**: teoría, metodología, aplicaciones. Barcelona: Empúria, 1993. p. 11-17. (Prólogo).

SMIT, J. W. A documentação e suas diferentes abordagens. In: GRANATO, M. (Org.). **Documentação em Museus**. Rio de Janeiro: MAST – Museu de Astronomia e Ciências Afins, 2008, v. 1, p. 11 – 22.

SMIT, J. W. A representação da imagem. **Informare**, Rio de Janeiro, v. 2, n. 2, p. 28 – 36, 1996.

SMIT, J. W. O arquivo de museu e a informação. In: SEMINÁRIO INTERNACIONAL ARQUIVOS DE MUSEUS E PESQUISA, 1., 2010. São Paulo. **Anais...** São Paulo: MAC USP, 2010, p. 84 – 92.

TÁLAMO, M. de F. G. M. Linguagem documentária. **Associação Paulista de Bibliotecários – Ensaio APB**, São Paulo, n. 45, 1997.

TÁLAMO, M. de F. G. M.; LARA, M. L. G. de. O campo da linguística documentária. **Transinformação**, Campinas, v. 18, n. 3, p. 203 – 211, set. / dez. 2006.

THESAURI& CONTROLLED VOCABULARIES - Library of Congress. Disponível em: <<http://www.loc.gov/library/libarch-thesauri.html>>.

VOGT, C. A espiral da cultura científica. **Revista Com Ciência**, n. 45. Jul. / 2003.
(Editorial).

WERSIG, G. Information Science: the study of postmodern knowledge usage. **Information Processing & Management**, v. 29, n. 2, p. 229 -239, 1993.

ZINS, C. Knowledge organization: an epistemological perspective. **Knowledge Organization**, v. 31, n. 1, p. 49 – 54, 2004.

VARIÁVEIS DA ANÁLISE DOCUMENTÁRIA DE CHARGE: UMA PROPOSTA INICIAL

VARIABLES IN THE SUBJECT ANALYSIS OF POLITICAL CARTOONS: AN INITIAL PROPOSAL

Thulio Pereira Dias Gomes
Rosali Fernandez de Souza

Resumo: Discute as variáveis da análise documentária aplicadas à charge. Justifica-se na intenção de contribuir para que a charge seja reconhecida como documento para a sua inserção em sistemas informativos-documentais. O objetivo é analisar aspectos endógenos e exógenos da charge para identificar elementos para representação documentária em sistemas de recuperação de informação. Pressupõe teorias de documento e informação no âmbito da análise documentária tendo como base o paradigma hermenêutico da informação relacionado à perspectiva domínio-analítica. Discute as variáveis da análise documentária de texto, propostas por Lara, considerando os seguintes elementos: o documento, a informação, a sistematização da área, a instituição, o usuário e o documentalista. Compara e interpreta esses elementos na análise documentária da charge. Como resultado, propõe as seguintes variáveis na análise documentária de charge: a) a estrutura discursiva da charge; b) as informações brutas presentes na charge; c) os domínios de atividade em torno da charge; d) instituições de custódia da charge; e) os usuários da charge e f) a formação e o quadro de referência do documentalista para o tratamento da charge. Essas variáveis evidenciam a circunscrição desse processo em condições sociais específicas. Por consequência, a análise de assunto da charge não pode ser assumida como neutra e imparcial, o que nos permite concluir sobre a inviabilidade de uma proposta universal e atemporal para análise documentária de charge.

Palavras-chave: Análise documentária. Charge. Variáveis da análise documentária.

Abstract: This paper discusses the variables in the subject analysis applied to political cartoon. It is justified by the intention to consider political cartoon as document in information retrieval systems. The objective is to analyze endogenous and exogenous features of political cartoon in order to identify elements for the representation of this kind of document in information retrieval systems. The theoretical framework considers the concepts of information and documents on the basis of the hermeneutic paradigm, domain analysis and subject analysis. Discusses the variables in the subject analysis of text, proposed by Lara, which considers the following elements: document, information, the systematization of the knowledge area, the institution, the user and the information professional. These elements are compared, interpreted and adapted in the subject analysis of political cartoon. As a result, it is proposed the following variables in the subject analysis of political cartoon: a) the discursive structure of political cartoon; b) the referential information in political cartoon; c) the domains around the political cartoon; d) custodial institutions of political cartoon; e) users of the political cartoon and f) information professionals profile and referral. These variables reflects the evidence of the circumscription process in specific social conditions. Thus, the subject analysis of the political cartoon cannot be neutral and impartial. This understanding led to the conclusion on the unfeasibility of a universal and non-temporal proposal for subject analysis of political cartoon.

Keywords: Subject analysis. Historical, political cartoon. Variables in the subject analysis.

1 INTRODUÇÃO

Este trabalho aborda as variáveis da análise documentária (AD) de charge. Justifica-se na intenção de contribuir para que a charge seja reconhecida como documento para a sua inserção em sistemas informativos-documentais. O objetivo é identificar, a partir das seis variáveis da análise documentária, propostas por Lara (1993), o que pode ser considerado como variável no contexto do tratamento temático de charge, documento de linguagem visual.

Entende-se por charge o “gênero discursivo de uso híbrido das linguagens verbal e imagética, caracterizado pela temporalidade marcada pela sátira e pela crítica referentes a determinado evento, em geral de natureza política” (GOMES, 2013, p. 26-27). A FIGURA 1 apresenta um exemplo de charge segundo a definição proposta.

FIGURA 1. Charge de Jean Galvão, para o jornal Folha de São Paulo, publicada em 19 de julho de 2013.



A charge de Jean Galvão foi publicada no contexto da visita do Papa Francisco ao Brasil, ocasionada pela Jornada Mundial da Juventude, evento católico ocorrido na capital fluminense, em julho de 2013. Na charge, duas figuras representam organizadores do evento que discutem a escolha do “papanóvel” por Sua Santidade. O papa optou pelo veículo sem vidros, possivelmente para transmitir a ideia de maior proximidade com os fiéis. O veículo blindado já teria sido pedido por “Sérgio Cabral”, o governador do Rio de Janeiro, quem na época estava sendo alvo de diversos protestos por todo o Estado e acusado de utilizar aeronaves do governo para fins particulares.

Na charge, a simultaneidade das linguagens imagética e verbal é notável no desenho e no texto da fala dos personagens nos balões. É importante observar a complementaridade das linguagens, porque a compreensão da mensagem só é possível se forem levados em conta os dois códigos. A temporalidade marcada é observada na referência a um evento específico, isto é, a onda de protestos que marcaram o governo de Sérgio Cabral na época da visita do papa ao

Brasil. A crítica e a sátira a esse evento são notáveis na marcação da opinião do chargista sobre o evento.

2 EM BUSCA DAS VARIÁVEIS DA ANÁLISE DOCUMENTÁRIA DE CHARGE

A análise documentária (AD) consiste em “uma atividade metodológica específica no interior da Documentação, que trata da análise, síntese e representação da informação, com o objetivo de recuperá-la e disseminá-la” (CINTRA et al., 1994, p. 24). Lara (1993) identifica no processo global da AD seis variáveis. Para a autora, este processo varia de acordo com:

a) a estrutura do texto original

A estrutura textual é o reflexo do arranjo utilizado pelo autor para a apresentação das informações (LARA, 1993). Para Mollica e Guedes (2013), os falantes embalam as informações por meio de opções linguísticas e de estruturas da linguagem em contextos de fala e de escrita. O reconhecimento dessas estruturas favorece a compreensão do fluxo da informação durante a leitura do texto.

Em termos de AD de charge, a dimensão textual, porém, não esgota uma das principais características da charge, a qual é a sua relação com o contexto de enunciação (FLÔRES, 2002; GOMES; GUEDES; SANTOS, 2013). Torna-se, portanto, necessária a transposição do nível textual para o discursivo. Os conhecimentos da estrutura do discurso e de como se dá a produção sófica podem permitir a elaboração de estratégias para a leitura documentária de charge. A transposição para o nível discursivo suscita a discussão sobre a proficiência da AD, em seus diálogos com a linguística textual de abordagem interacionista, para o tratamento da charge. Tal corrente teórico-metodológica se mantém válida na medida em que enfatiza os procedimentos documentários, porém é necessária interlocução com outras teorias que considerem a dimensão discursiva da charge.

A estrutura textual traz conflitos entre a linguagem da charge e a do sistema informativo documentário. Segundo Foucault (2000, p. 11), “por mais que se diga o que se vê, o que se vê não se aloja no que se diz”. Para o filósofo, o lugar das palavras não é o lugar descortinado pelos olhos, mas é um lugar definido pelas sucessões da sintaxe. Com base na afirmativa de Foucault, é possível declarar a impossibilidade de uma efetiva tradução de conteúdos da charge, e de qualquer outra imagem, para as linguagens documentárias. Os resumos, as legendas e os descritores não são eficazes para a representação da totalidade do conteúdo da charge, tampouco das possibilidades de significação e de utilização. De fato, “analisar uma imagem significa, quer queiramos quer não, ‘traduzir’ certos elementos desta imagem de um código icônico para um código verbal” (SMIT, 1989, p. 105). Sob outra

perspectiva, também é possível questionar como os usuários podem traduzir suas necessidades de informação, em vista das restrições da linguagem do sistema.

b) a informação bruta presente no texto

As informações brutas presentes no texto, conforme Lara (1993), se referem às mensagens mais significativas. Segundo Vanoye (2002), as informações brutas são integradas por elementos referenciais, ou seja, elementos que remetem explicitamente ao espaço e ao tempo em que a mensagem é veiculada. Dessa forma, as informações brutas e objetivas no texto configuram uma das variáveis do processo de AD da charge. Entretanto, a leitura da charge pode falhar na identificação de conteúdos informativos, conforme os estudos de Gomes, Guedes e Santos (2013) e de Wu (2013). São necessários recursos teóricos e metodológicos que favoreçam a identificação desses elementos referenciais da charge, isto é, as informações de base contidas na charge, as quais são condições de sua informatividade.

No âmbito do paradigma hermenêutico da CI, relacionado à perspectiva domínio-analítica, a informatividade depende da relação entre a relevância da mensagem comunicada e do conhecimento prévio de certo indivíduo e/ou do conhecimento compartilhado por determinada comunidade discursiva. Para haver informação, é necessária uma pré-estrutura de conhecimento capaz de contextualizar e de interpretar a mensagem, gerando uma nova estrutura de conhecimento. A informação, assim, se constitui no conhecimento em prática e está relacionada à novidade e à relevância do conteúdo das mensagens (CAPURRO; HJØRLAND, 2003).

Recorre-se também à noção de informação na Linguística Documentária. Segundo a proposta de Lara (2006 *apud* LARA, 2008), a informação é o “resultado de uma construção institucional e intencional que tem nos valores simbólicos e funcionais a condição para a construção do sentido e para circular socialmente, desencadeando processos de conhecimento”. A construção de sentido e a circulação social, condicionadas aos valores simbólicos e funcionais, propõem a ideia de que a informação não é uma questão individual, mas de ordem social e cultural.

Uma das principais características da charge é a referência a outros eventos e textos, cabendo ao leitor estabelecer elos para construção do sentido proposto pelo chargista. Se a interação entre autor, texto e leitor não se estabelece, as “informações contidas na charge não se convertem em signos, não são percebidas e, nesse caso, não entram no circuito comunicacional” (FLÔRES, 2002, p. 11). O interdiscurso e a intertextualidade caracterizadores da charge a aproximam da ideia de informatividade de Capurro e Hjørland (2007) e do entendimento de informação como um processo construtivo sugerido por Lara

(2008). Portanto, se ocorrer efetiva compreensão pelo leitor, a leitura da charge se torna um processo informativo em que um usuário possui uma estrutura prévia de conhecimento a partir da qual interpretará a mensagem, isto é, identificará a informatividade do documento segundo seus próprios critérios de relevância.

c) o estado de sistematização da área em questão

Entre as evidências da sistematização das áreas do conhecimento está a sua terminologia. A AD é facilitada em áreas onde há maior precisão terminológica e dificultada naquelas onde há precisão menor. A imprecisão terminológica dos domínios para os quais se representa a charge pode ser apontada como uma dimensão desta variável. A variação no emprego do termo charge na linguagem corrente e na literatura especializada é um exemplo de configuração dessa variável no âmbito das linguagens desenhadas.

d) a instituição na qual o processo se desenvolve

O contexto geral dentro do qual a AD se insere é denominado por Lara (1993) como instituição. Fujita (2004) estende a noção de contexto da AD indicando uma conjunção de elementos como os objetivos da indexação, a política e o manual de indexação, a linguagem documentária, o próprio indexador e o usuário. Esses elementos, por sua vez, variam conforme a missão, a visão, os objetivos e recursos de cada uma das instituições.

As políticas e os manuais de indexação estabelecem procedimentos da representação documentária, interferindo em todas as fases da recuperação de documentos no sistema. A diversidade de propostas de análise de imagens encontradas na literatura salienta as tantas possibilidades de fundamentos teórico-metodológicos para a AD de charge. Tal diversidade sugere adaptações ao contexto, evidências da influência da instituição sobre a AD da charge.

e) o usuário da informação documentária

A instituição delimita o usuário e estabelece seu perfil (LARA, 1993). Se analisada desde a noção de domínio, essa variável tem algumas consequências. Domínio pode ser caracterizado como um grupo de pessoas que compartilham objetivos, como, por exemplo, uma área de especialidade, um campo da literatura ou um grupo de pessoas trabalhando juntas numa organização (MAI, 2005). O trabalho de um domínio reflete sobre a organização do conhecimento, os padrões de cooperação, a linguagem e os meios de comunicação, nos sistemas de informação e os critérios de relevância (HJØRLAND; ALBRECHTSEN, 1995).

A inserção da charge em um sistema informativo documental deve considerar os diversos usos da charge pelos diferentes domínios. Cada domínio utiliza a charge segundo seus próprios critérios de relevância. Apesar das evidências de uso da charge, não há análises de domínio caracterizadoras da comunidade, ou das comunidades, em torno da

charge. A definição do domínio em torno da charge permite a identificação da linguagem adequada para ser utilizada na representação documentária bem como dos critérios de relevância de diferentes grupos servidos pelo sistema informativo documental.

f) perfil do documentalista

A formação e o quadro de referência do documentalista determinam os processos de AD (LARA, 1993). Para Fujita (2004), o objetivo do documentalista é a representação do assunto de um documento por meio de termos significativos, para que seja possível a recuperação pelos usuários interessados. O êxito do processo depende das habilidades e dos conhecimentos prévios do profissional, necessários à atividade documentária.

A leitura da charge é um processo informativo de contextualizar o conhecimento objetivado no documento, a partir de um pressuposto já conhecido. Quanto maior o conhecimento pressuposto, maior a chance do processo informativo durante a AD alcançar o sucesso. Esta discussão suscita questões sobre quais os conhecimentos e as competências o documentalista deve possuir para tratar a charge em dado sistema informativo documental.

Cunha (1990, p. 71) discute o problema da ideologia do documentalista na execução das atividades profissionais. Para a autora, a negação do componente ideológico em análise do texto para fins documentais “é não colocar em questão o papel do indivíduo no processo de produção, consumo e tradução do conteúdo para o todo/conceito/palavra-chave”. A autora alerta que essa postura consiste em negar que a linguagem do discurso ou da linguagem documentária é aquisição, construção e instrumento “elaborado respectivamente pela sociedade por e para um grupo social, refletindo, em ambos os casos, o pensamento e os interesses dominantes”. Tais questões fomentam discussões a respeito das limitações de tal processo considerando as diferenças entre os contextos ideológicos de produção e de leitura.

3 CONCLUSÃO

Em função da discussão acima, é possível propor a identificação das seguintes variáveis da análise documentária de charge: a) a estrutura discursiva da charge; b) as informações brutas presentes na charge; c) os domínios de atividade em torno da charge; d) instituições de custódia da charge; e) os usuários da charge e f) a formação e o quadro de referência do documentalista para o tratamento da charge.

A existência de variáveis da AD é uma evidência de que esse processo de organização do conhecimento é histórico, ou seja, circunscrito em condições sociais específicas. Por consequência, a representação documentária da charge não pode ser assumida como neutra e

imparcial, o que nos permite concluir sobre a inviabilidade de uma proposta universal e atemporal para análise documentária de charge.

REFERÊNCIAS

- CAPURRO, R.; HJØRLAND, B. O conceito de informação. **Perspectivas em Ciência da Informação**, Belo Horizonte, MG, v. 12, n. 1, p. 148-207, jan./abr. 2007.
- CINTRA, A. M. M. et al. **Para entender as linguagens documentárias**. São Paulo: Editora Polis, 1994.
- CUNHA, I. M. R. F. **Do mito à análise documentária**. São Paulo: EDUSP, 1990.
- FLÔRES, O. **A leitura da charge**. Canoas, RS: Ed. ULBRA, 2002.
- FOUCAULT, M. **As palavras e as coisas**. São Paulo: Martins Fontes, 2000.
- FUJITA, M. S. L. A leitura documentária na perspectiva de suas variáveis. **Datagramazero**, v. 5, n. 4, Artigo 01, ago. 2004.
- GOMES, T. P. D. **Temas e questões em análise documentária de charge**. 2013. 65 f. Trabalho de Conclusão de Curso–(Graduação em Biblioteconomia). Universidade Federal do Rio de Janeiro. Rio de Janeiro, 2013.
- GOMES, T. P. D.; GUEDES, V. L. da S.; SANTOS, M. J. V. da C. Interferências dos contextos de produção e de uso na análise documentária de charge. In: RIBEIRO, F.; CERVEIRA, M. E. (Org.). **Informação e/ou Conhecimento**. Porto: FLUP, 2013. p. 245-263.
- HJØRLAND, B.; ALBRECHTSEN, H. Toward a new horizon in information science. **Journal of the American Society of Information Science**, v.46, n. 6, p. 400-425, 1995.
- LARA, M. L. G. de. Informação, informatividade e Linguística Documentária. **Datagramazero**, v. 9, n. 6, Artigo 01, dez. 2008.
- LARA, M. L. G. **A representação documentária**. 1993. 133 f. Dissertação (Mestrado em Ciências da Comunicação)– Universidade de São Paulo, São Paulo, 1993.
- MAI, J.-K. Analysis in indexing. **Information Processing and Management**, v. 41, p. 599-611, 2005.
- MOLLICA, M. C. de M.; GUEDES, V. L. da S. Mudança no fluxo informacional e gêneros discursivos. In: ALBAGLI, S. (Org.). **Fronteiras da Ciência da Informação**. Brasília, DF: IBICT, 2013. p. 216-235.
- SMIT, J. W. A análise da imagem. In: GRUPO TEMMA. **Análise documentária**. 2. ed. Brasília, DF: IBICT, 1989. p. 102-113.
- VANOYE, F. **Usos da linguagem**. São Paulo: Martins Fontes, 2002.
- WU, Y. Indexing historical, political cartoons for retrieval. **Knowledge Organization**, v. 40, n. 3, p. 283-294, 2013.

NANOPUBLICAÇÕES E INDEXAÇÃO: ASPECTOS METODOLÓGICOS

NANOPUBLICATIONS AND INDEXING: METHODOLOGICAL ASPECTS

Lorena Tavares de Paula
Maria Aparecida Moura

Resumo: As nanopublicações são consideradas “a menor unidade de informação publicável”. Esta menor unidade publicável refere-se a conceituação de termos representativos de um determinado domínio de conhecimento. Para efetuar a representação das nanopublicações, para fins de recuperação da Informação, foi proposta uma investigação que evidenciasse a interlocução entre os processos de indexação e a modelagem das nanopublicações. Nesse artigo apresenta-se e contextualiza-se o conceito de nanopublicação e uma sugestão metodológica para a sua modelagem auxiliada por princípios da indexação. Apresentam-se ainda considerações sobre métodos e conceitos que podem auxiliar na consolidação desse modelo emergente de representação de informação. Salienta-se ainda, que as nanopublicações surgem como um desafio a ser enfrentado pelo campo da Ciência da Informação. Tudo do ponto de vista das mudanças nas formas de produção e circulação da informação e do conhecimento, as quais enunciam a exigência de articulação de metodologias adequadas para dar visibilidade e acesso a esses fluxos intensos de informação. Concluiu-se que, produzir conhecimento de maneira ágil exige mecanismos de recuperação de informação tão precisos quanto as unidades citáveis caracterizadas pelas nanopublicações.

Palavras-chave: Indexação, Nanopublicação, Metodologia, Cibercultura.

Abstract: Nanopublications are considered to be “the smallest publishable units of information.” This smallest publishable unit refers to the concept of representative terms of a particular domain of knowledge. In order to represent nanopublications, for recovery purposes of information, we proposed an investigation that showed the dialogue between processes of indexing and modeling of nanopublications. In this paper it is presented and contextualized the concept of nanopublication and a methodological suggestion for its aided modeling principles of indexing. It is still presented considerations on methods and concepts that can help to consolidate this emerging model of representation of information. It is also noted that nanopublications arise as a challenge to be faced by the field of Information Science. Everything from the point of view of changes in production and circulation of information and knowledge forms set out the requirement to articulate appropriate methodologies to provide visibility and access to these intense information flows. Producing knowledge in an agile manner requires mechanisms for information retrieval as accurate as the citable units characterized by nanopublications.

Keywords: Indexation, Nanopublication, methodology, Cyberculture

1 INTRODUÇÃO

Um dos destaques, dentre as inovações na Internet, é a possibilidade de acesso há uma considerável quantidade de informações organizadas e armazenadas pelo próprio usuário em um ambiente da internet para posterior utilização em contextos e geografias distintas. Nestes ambientes as técnicas e tecnologias orientadas à organização da informação buscam promover uma melhor relação entre usuários e conteúdos considerando as possibilidades da

infraestrutura digital e suas alternativas para a colaboração e disseminação informacional nas “nuvens”.

Observando o usuário como gestor de seus conteúdos, há de se estabelecer estratégias de recuperação da informação que ofereçam agilidade e precisão em suas interações com as informações depositadas. Neste sentido, o presente artigo pretende apresentar e discutir as nanopublicações, que são consideradas a menor unidade de informação publicável, a fim de estabelecer uma interface com o processo de indexação de informação, objetivando:

1. Relacionar nanopublicações e indexação
2. Demonstrar a possibilidade de indexação no texto completo a partir das nanopublicações
3. Estabelecer uma metodologia que oriente uma indexação intuitiva de conteúdos relevantes aos usuários.

Para tanto serão expostas definições conceituais pertinentes aos assuntos abordados neste artigo para em sequência estabelecer uma sugestão metodológica para a construção de nanopublicações que integre os processos da indexação como princípios norteadores para a sua modelagem. Por fim, apresentam-se as considerações sobre a compreensão das nanopublicações como mecanismo emergente de representação da informação.

2 NANOPUBLICAÇÃO

A nanopublicação pode ser considerada a menor unidade de informação publicada em um documento formal: trata-se de uma extração textual atribuída a “fala do autor”.

Segundo a *Concept Web Alliance (CWA)*, (2013, *online*) as nanopublicações podem ser citadas, pois são uma referência exata das ideias contidas em uma publicação formal. Elas possibilitam uma comunicação que se baseia em informações contextuais de alta qualidade e que servem para a divulgação, apropriação e organização de informações contextualizadas de acordo com um autor e uma unidade conceitual.

Groth *et al.* (2011, *online*) explicam que, em seu aspecto conceitual as nanopublicações devem ser citáveis, atribuíveis e revisáveis. Além disso, elas precisam ser reorganizáveis sem dificuldades, pois “as nanopublicações podem ser facilmente agregadas e identificadas em toda a web”. Elas precisam ser extensíveis para atender as novas formas de descrição orientadas por conjuntos de definições como:

1. Definição Conceitual: conceito entendido como a menor unidade, livre de ambiguidades;
2. Sintaxe: envolvimento de três conceitos (sujeito, predicado, objeto);

3. Declaração – Uma sintaxe que é exclusivamente identificável;
4. Anotação – Uma declaração sobre a sintaxe que reforce o assunto e seja referenciável;

Groth *et al.* (2011, *online*) esclarecem que a nanopublicação pode ser considerada um conjunto de anotações que se referem a declarações e conteúdos orientados a comunidades que propõem as definições conceituais. A partir disso, ela pode atender a alguns requisitos básicos como:

- A capacidade de identificação conceitual;
- A capacidade de identificação de comunidades;
- Capacidade de compilação de conceitos e declarações;
- Em uma perspectiva computacional, permitir interoperabilidade com diferentes formatos computacionais.

Diante da sua configuração computacional ocorrem também os requisitos

- *fromSource* (identifica onde a declaração de pesquisa foi extraída);
- *importedBy* (identifica que a entidade é responsável pela importação da declaração);
- *authoredBy* (identifica o autor de uma declaração de pesquisa).

Observa-se que essa modalidade de publicação faculta aos autores, editores e pesquisadores a possibilidade de disponibilizarem seus conteúdos em formatos que impulsionam o acesso de suas produções de acordo com o que consideram mais relevantes para o impacto sócio/comunicacional de sua obra. Deste modo, as nanopublicações apresentam dois elementos básicos:

1. A afirmação: Uma afirmação é uma unidade mínima do pensamento, que expressa uma relação entre dois elementos (chamados o sujeito e o objeto) usando um terceiro elemento (chamado de predicado).
2. A Proveniência: Este elemento fornece o contexto sobre a afirmação. Proveniência significa, “como isso veio a ser” e pode incluir a descrição do método e algumas atribuições (como autores, instituições, tempo, subvenções, *links* internos e externos).

Ressalta-se que as nanopublicações podem ser realizadas usando ontologias existentes e *RDF* (*Resource Description Framework*) e permitem a leitura em processo de interoperabilidade. Além disso, há a possibilidade de compatibilizar informações heterogêneas e descentralizadas, contribuindo para a descoberta de novas associações, que auxiliam na construção de novos conhecimentos.

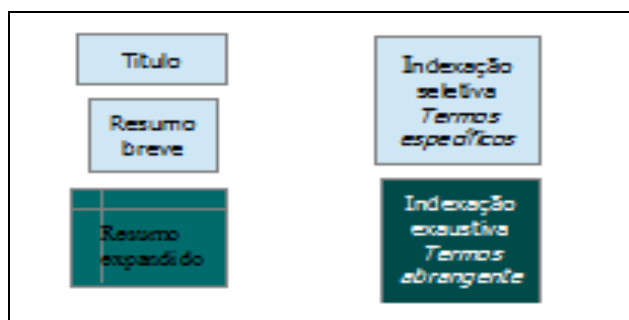
3 PROCESSOS DE INDEXAÇÃO

O processo de indexação obedece a dois princípios básicos: análise conceitual; tradução para uma linguagem de indexação. Lancaster (2004, p. 27) explica que a indexação e o resumo são “atividades intimamente relacionadas, pois ambas implicam a preparação de uma representação do conteúdo temático dos documentos”. O indexador emprega vários termos de indexação que possuem relação com o conteúdo intelectual do documento e a provável indagação de busca de informação do usuário a fim de estabelecer uma interação eficiente entre sujeito e mecanismo de recuperação de informação.

Uma indexação de assunto eficiente implica que se tome uma decisão não somente quanto ao que é tratado no documento, mas também por que ele se reveste de provável interesse para determinado grupo de usuários.

O processo de análise pode observar as seguintes estruturas no processo de indexação.

FIGURA 1 - Descrição do Registro



Fonte: elaborado pela autora, baseado em Lancaster (2004).

A FIG. 1 é um modelo baseado no “efeito do registro sobre recuperabilidade” de Lancaster (2004, p. 8), percebe-se que quanto mais informações são apresentadas, mais claramente a representação revela o alcance do documento. Deste modo, o leitor poderá ter o máximo de certeza se essa informação satisfaz ou não suas necessidades. A princípio observa-se o “título”, que pode não oferecer especificidade sobre os tópicos tratados no documento. No “resumo breve”, pode-se obter informações focalizadas no tópico central da publicação. Já o “resumo ampliado” oferece conteúdos abrangentes que correspondem às várias seções da publicação, bem como algumas de suas intenções. A “indexação exaustiva” irá incluir termos que indicam de maneira geral todos os assuntos plausíveis de indexação no documento e a “indexação seletiva” proporcionará uma indicação minuciosa dos assuntos específicos e relevantes.

Hjørland (2001, p.776) acrescenta que;

Uma vez que qualquer documento pode, em princípio, proporcionar respostas a uma infinidade de questões, as análises de assunto devem

estabelecer propriedades baseadas nos grupos de usuários específicos atendidos. O assunto de um documento é assim relativo ao objetivo do serviço de informação específico. Defino assunto [...] como potenciais epistemológicos ou informativos dos documentos. A melhor análise temática é a que faz o melhor prognóstico quanto ao uso futuro do documento.

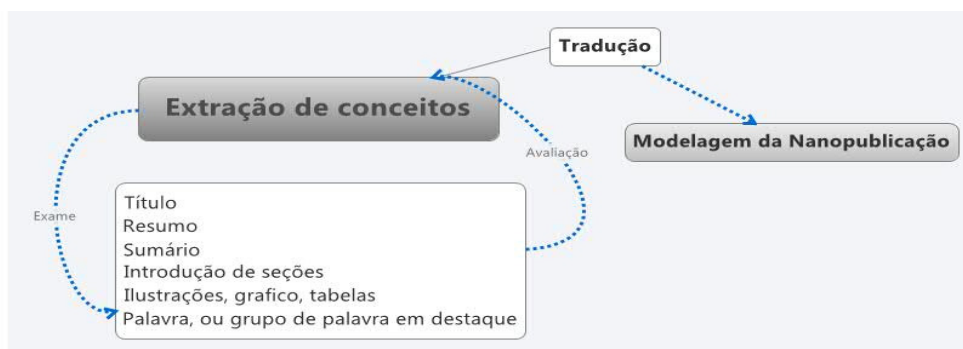
Todo o processo de indexação objetiva exclusivamente a recuperação da informação. Salienta-se que a análise de assunto é a chave para toda a indexação efetiva. Neste processo, é sempre importante analisar o documento em suas minúcias e estabelecer uma análise tão exaustiva quanto a relevância do documento para a comunidade de usuários da unidade de informação. Neste sentido, a indexação e as nanopublicações convergem com a necessidade de identificação de conteúdos que ofereçam especificidade ao documento no momento da recuperação da informação.

4 NANOPUBLICAÇÃO E INDEXAÇÃO: SUGESTÃO METODOLÓGICA

As orientações metodológicas descritas nesta seção têm a finalidade de elucidar os objetivos apresentados na introdução desse artigo, os quais almejam: relacionar as nanopublicações e a indexação; demonstrar a modalidade de indexação de fragmentos conceituais em texto completo. Também pretende-se estabelecer uma metodologia que oriente uma indexação intuitiva de conteúdos relevantes para o usuário em seu acervo particular de modo que o documento possa ser identificado e compartilhado com outros sujeitos na internet.

Para tanto, inicialmente serão delimitados 2 (dois) processos analíticos para extração e assimilação de informações de publicações, o primeiro baseado nas premissas da indexação e o segundo relacionado a junção: sujeito, predicado e objeto, característica das nanopublicações.

FIGURA 2 - Processo de extração de conceitos

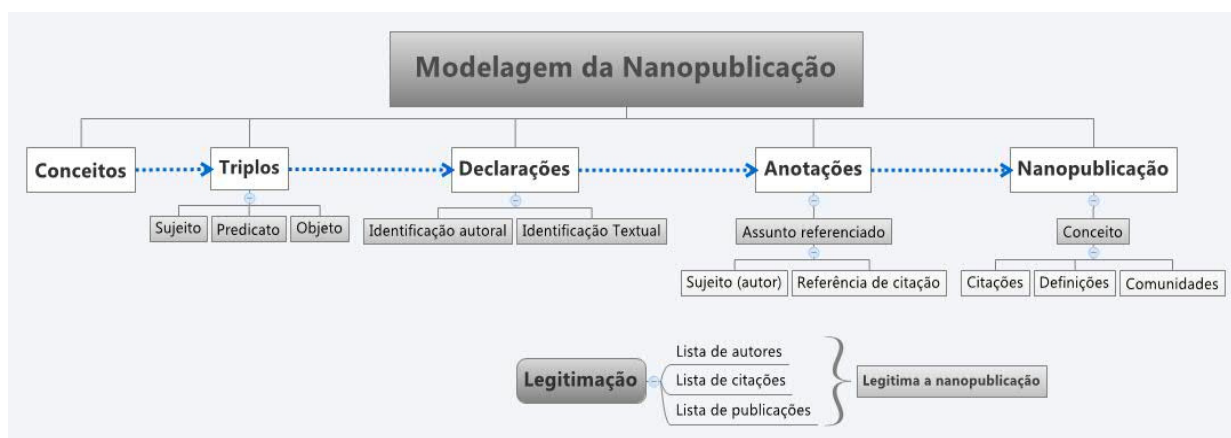


Fonte: elaborado pelas autoras, baseado em Lancaster (2004).

Na FIG. 2, percebem-se os principais elementos textuais, nos quais a extração de conceitos pode ser efetivada: título, resumo sumário, introduções de seções, ilustrações e palavras em evidência. O destaque para esses elementos não quer dizer que a extração de conceito deve se limitar a eles, mas que esses são espaços em que normalmente se encontram os insumos necessários para identificação e/ou localização de definições conceituais relevantes. Além disso, o conhecimento prévio sobre o domínio no qual se realiza a extração de conceitos faz-se imprescindível para uma verificação crítica das opções adotadas. Durante a fase de extração conceitual não há necessidade de consulta a um vocabulário controlado, esse exercício de tradução será pensado na etapa de elaboração da nanopublicação.

Feita a identificação de ocorrência do conceito central a ser extraído, faz-se em sequência um exame da definição conceitual estabelecida, para posteriormente estabelecer uma avaliação das definições que irão compor a nanopublicação, momento em que serão propostas relações entre os termos, a fim de compor as declarações e anotações que legitimam essa construção representacional.

FIGURA 3 – Modelagem da nanopublicação



Fonte: elaborado pelas autoras (2014).

A FIG. 3 apresenta o processo de modelagem de uma nanopublicação. De início são estabelecidos os conceitos que foram retirados durante a fase de análise de assunto baseada nos princípios de indexação. Estes conceitos são considerados a menor unidade sem ambiguidades que expressa um pensamento exclusivo e identificável. Posteriormente, são ajustados os processos de triplo (syntaxe), no qual são situadas as relações entre: sujeito, predicado e objeto, neste momento ocorre a relação de ação (causa/efeito) sobre o conceito e inicia-se a elaboração da nanopublicação.

A construção da declaração é estabelecida a partir das relações entre os termos extraídos, quando pode-se elaborar um mapa conceitual, e a identificação das definições

extraídas dos artigos com suas identidades autorais. As anotações veem em sequência para oferecer uma declaração que abranja o assunto estabelecido ao longo da modelagem. Por fim, a nanopublicação é estabelecida como um conjunto de anotações que consolidam declarações de publicações oriundas de comunidades distintas. Além disso, são estabelecidas as evidências, no qual são ratificadas as sugestões de metadados, para a representação do “conceito” como “menor unidade citável”. Deste modo pode-se proporcionar confirmações bibliográficas de todas as afirmações ajustadas na nanopublicação. Em resumo, no momento da especificação das e-evidências são acordados os metadados apropriados para a identificação da nanopublicação, como por exemplo: comunidades de produção do conceito; autores canônicos; unidade de citação, que refere-se a extração textual mais citada, etc.

5 CONSIDERAÇÕES FINAIS

A intensa adoção de recursos tecnológicos nos processos de produção do conhecimento promove constantes inovações nos mecanismos de identificação, representação e disseminação de informações. A obsolescência tem se caracterizado em segmentos que vão desde a elaboração de *hardwares* até os mecanismos (softwares) de produção e distribuição de conteúdos em meio digital. Neste contexto, a elaboração de nanopublicações faz-se eficiente para potencializar o compartilhamento de dados com a finalidade de gerar novos conhecimentos.

Diante dos visíveis avanços do processo de produção e difusão do conhecimento tornou-se necessário buscar metodologias que impulsionem o uso da crescente produção documental, através de mecanismos de representação de informações coerentes com as demandas sócio-tecnológicas vigentes. Produzir conhecimento de maneira ágil, como normalmente é proposto pelos órgãos de fomento à pesquisa em Ciência e Tecnologia, tende a exigir mecanismos de recuperação de informação tão precisos quanto as unidades citáveis caracterizadas pelas nanopublicações.

Nesse sentido, as nanopublicações surgem com um desafio a ser enfrentado pelo campo da Ciência da Informação. Tudo do ponto de vista das mudanças nas formas de produção e circulação da informação e do conhecimento, as quais enunciam a exigência de articulação de metodologias adequadas para dar visibilidade e acesso a esses fluxos intensos de informação.

REFERÊNCIAS

CHICHESTE, C. et al. **The Open PHACTS Nanopublication Guidelines**, [S.l], [s.n.], 2012. Disponível em: <<http://www.nanopub.org/guidelines/1.8/>>. Acesso em: 14 mar. 2014.

CONCEPT WEB ALLIANCE (CWA). **Nanopublication Guidelines**, Working Draft, [S.l], 2013. Disponível em: <http://nanopub.org/guidelines/working_draft/>. Acesso em: 8 set. 2014.

GROTH, P. *et al.* The Anatomy of a Nano-publication, *Information Services & Use* **V.30**, [S.l], 2011. Disponível em: <[http://www.w3.org/wiki/images/c/c0/hclsig\\$\\$swansioc\\$\\$actions\\$\\$rhetoricalstructure\\$\\$meetings\\$\\$20100215\\$cwa-anatomy-nanopub-v3.pdf](http://www.w3.org/wiki/images/c/c0/hclsig$$swansioc$$actions$$rhetoricalstructure$$meetings$$20100215$cwa-anatomy-nanopub-v3.pdf)>. Acesso em: 14 mar. 2014.

HJORLAND, B. Towards a theory of aboutness, subject, topicality, theme, domain, field, content and relevance. **Journal of the American Society for Information Science and Technology**, n. 52, p. 774-778, 2001.

LANCASTER, F. W. **Indexação e resumo**: teoria e prática. Brasília: Briquet de Lemos/Livros, 2004.

MONS, B; VELTEROP, J. Nanopublication in the e-science Era. In: WORKSHOP ON SEMANTIC WEB APPLICATIONS IN SCIENTIFIC DISCOURSE, INTERNATIONAL SEMANTIC WEB CONFERENCE, 8., 2009, Washington, DC, 2009.

CATEGORIZAÇÃO DE PATENTES COM USO DE INTELIGÊNCIA COMPUTACIONAL

CATEGORIZATION OF PATENTS USING COMPUTATIONAL INTELLIGENCE

Magali Rezende Gouvêa Meireles
Beatriz Valadares Cendón

Resumo: As patentes são uma fonte de informação sobre a criatividade humana focada em conhecimento prático aplicado e, por isso, tornam-se indicadores importantes para se medir o avanço tecnológico de um domínio específico de conhecimento. Porém, o volume de patentes disponíveis em bancos digitais desafia o usuário que busca informações que possam lhe ser úteis em processos de tomada de decisão relacionados à identificação de áreas que estão em significativo desenvolvimento e áreas que apresentam boas oportunidades de inovação por conterem poucas pesquisas em seus campos. Este projeto tem como objetivo criar um sistema de categorização alternativo, que possibilite acesso simplificado às informações incorporadas nas patentes, e considera como hipótese principal a viabilidade de se extrair informações que podem gerar conhecimento em processos de decisão relativos à identificação de possibilidades de investimento de pesquisa. A partir de diferentes abordagens existentes na literatura para categorizar patentes, pretende-se avaliar os métodos existentes e propor, a partir de bases digitais de patentes disponíveis, um método de categorização adequado a este propósito. Por não existir um consenso entre os especialistas de que exista um melhor e completo processo de categorização e de classificação de patentes, não havendo padronização destes processos, serão investigadas métricas que possam ser utilizadas na avaliação destes sistemas. Serão utilizadas técnicas computacionais, tomando-se como atributo o conteúdo, os metadados e as citações das patentes. Do ponto de vista científico, o desenvolvimento desta pesquisa e a implementação de um processo alternativo de categorização de patentes configuram a proposta de um projeto interdisciplinar que utiliza os conceitos relacionados às tecnologias da informação e as ferramentas e conceitos de Inteligência Computacional.

Palavras-chave: Categorização. Inteligência Computacional. Patentes. Recuperação de Informação.

Abstract: Patents are an information source on human creativity focused on applied practical knowledge and, therefore, become important indicators to measure the technological advancement of a specific domain of knowledge. However, the volume of patents available in digital bases challenges the user who seeks information that may prove useful in decision-making processes relating to the identification of areas that are in significant development and areas that present good opportunities for innovation for having small amount of research on their fields. This project aims to create an alternative system of categorization, which allows easier access to the information embedded in patents, and considers as the main hypothesis the feasibility to extract information that can generate knowledge in the decision making processes related to opportunities for research investment. Starting from different approaches in the literature to categorize patents, we intend to evaluate existing methods and to propose, from available digital databases of patents, a suitable method of categorization. Because there is not a consensus among experts that there is a good and complete process of categorization and classification of patents and no standardization of these processes, the creation of metrics will be investigated which can be used in the evaluation of these systems. Computational techniques will be used taking as attribute content, metadata and citations of patents. From a scientific point of view, the development of this research and the implementation of an alternative process of patents categorization constitute the proposal of an interdisciplinary

project that uses the concepts related to information technologies and tools and concepts of Computational Intelligence.

Keywords: Categorization. Computational Intelligence. Patents. Information Retrieval.

1 INTRODUÇÃO

Um campo, segundo Saracevic (1996, p.41), é definido pelos problemas discutidos e pelos métodos que utiliza para solucioná-los. Desta maneira, três são as características que constituem a razão da existência e da evolução da Ciência da Informação (CI). A primeira delas é a sua natureza interdisciplinar. As relações com a Ciência da Computação, com a Inteligência Artificial e com a Ciência Cognitiva estão se tornando mais visíveis em aplicações. A segunda característica refere-se à sua ligação com a tecnologia da informação. E a terceira característica ressalta o fato da CI, por ser caracterizada por uma forte dimensão social e humana, ser uma participante ativa na evolução da sociedade da informação.

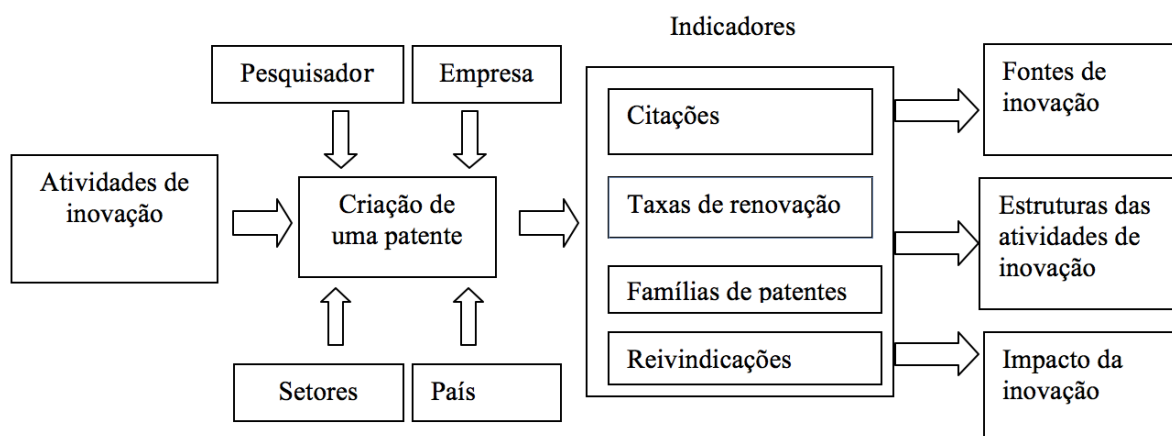
A CI vem dedicando grande interesse, em seus trabalhos de pesquisa, à solução ou melhoria de aspectos relacionados ao problema de organização e recuperação de informação. A recuperação da informação é definida por Abadal e Codina (2005, p.29) como a disciplina que estuda a representação, a organização e o acesso eficiente à informação registrada em documentos. As operações associadas à recuperação de informação têm como objetivo solucionar os problemas que requerem informação científica, técnica e cultural e que surgem em, pelo menos, alguma etapa da vida de todos os seres humanos, na tentativa de descobrir, estudar, aprender ou investigar um tema. Com o crescimento do volume de publicações disponível em meio digital, muitas técnicas de recuperação de informação foram desenvolvidas para atender as necessidades dos usuários e muitos estudos estão sendo realizados com o objetivo de organizar as coleções de documentos. As possibilidades de acesso às bases de dados disponíveis eletronicamente permitiram aumentar a qualidade das buscas bibliográficas e selecionar os documentos que atendem a necessidade da busca vem se tornando um processo desafiador para o usuário. Organizar as publicações para, posteriormente, recuperá-las é um tentativa de manter a qualidade da informação recuperada sem que possa ser comprometida pela quantidade cada vez maior de documentos disponíveis (MEIRELES, 2012, p.35).

A medida do avanço tecnológico de um domínio específico do conhecimento é, hoje em dia, uma importante proposta de pesquisa, que pode identificar caminhos inexplorados para se investigar. Informações detalhadas sobre o avanço da tecnologia são necessárias nos processos de tomada de decisão. Tais informações descrevem a quantidade de recursos destinados à inovação, a seleção de campos em que a inovação promete retornos econômicos

e a gestão de estratégias inovadoras (ARCHIBUGI; PIANTA, 1996, p.451). Camus e Brancaleon (2003, p.155) destacaram a importância da informação contida na análise de patentes, revelando os riscos e as oportunidades e obtendo conhecimento sobre atividades das empresas. Alguns estudos (MARKELLOS *et al.*, 2002; LEYDESDORFF, 2004) exploram as bases de dados de patentes, mostrando como a produção do conhecimento científico pode estar relacionada com a economia. Lee e outros (2009) afirmaram que as patentes são fontes completas de conhecimento técnico e comercial. Assim, a análise de seus atributos técnicos e de mercado tem sido considerada como uma ferramenta útil para pesquisa e desenvolvimento gerencial.

A FIG. 1 representa a atividade de criar uma patente como um indicador do desenvolvimento de um campo de tecnologia.

FIGURA 1 - A atividade de patentes como um indicador do desenvolvimento de um campo de tecnologia



Fonte: Adaptado de Archibugi e Pianta (1996).

Para utilizar as patentes como um indicador de desenvolvimento, é importante que as patentes estejam organizadas em grupos. O processo de categorização é um processo natural do ser humano, que busca, na criação de grupos, um modo de organizar as informações e simplificar suas interações com o mundo. As informações agrupadas se tornam mais facilmente entendidas, viabilizando e facilitando a interpretação e o entendimento das situações que cercam o indivíduo. Na literatura, podem ser encontradas algumas ferramentas de Inteligência Artificial utilizadas em processos de categorização e de classificação. A capacidade do cérebro de armazenar conhecimento e reorganizar as conexões já estabelecidas vem inspirando pesquisadores a aprimorar o desenvolvimento de uma delas, as Redes Neurais Artificiais (RNA).

A seção dois apresenta o referencial teórico do trabalho, ressaltando os conceitos utilizados na apresentação da proposta. As seções seguintes apresentam a metodologia utilizada e as considerações finais.

2 REFERENCIAL TEÓRICO

A fundamentação teórica deste projeto é dividida em três seções. A primeira seção discute os processos de categorização e de classificação. A segunda e a terceira seções apresentam conceitos relacionados a patentes e à Inteligência Computacional.

2.1 A Categorização e a Classificação

Segundo Jacob (2004), a categorização é o processo de dividir o mundo em grupos de entidades cujos membros têm similaridades entre eles dentro de um determinado contexto. Agregar as entidades em categorias leva o indivíduo a perceber ordem no mundo que o circunda. A classificação é uma capacidade conceitual humana, que utiliza as categorias como ferramentas. Para Jacob (2004), o processo de classificação envolve a associação ordenada e sistemática de cada entidade a uma única classe dentro de um conjunto de classes mutuamente exclusivas e que não se sobrepõem. É um processo que estabelece ordem ao demonstrar relações entre entidades, tanto explicitamente por meio das definições utilizadas para associar os objetos às classes, quanto implicitamente ao se construir uma estrutura com um relacionamento hierárquico entre as classes.

O processo de classificar é realizado de acordo com um conjunto de princípios estabelecidos que governa a estrutura e o relacionamento das classes. Estes princípios são aplicados sistematicamente na busca de uma reprodução ordenada do mundo real. Segundo Jacob (2004), o processo é artificial, porque é criado com o propósito de estabelecer uma organização significativa, e é arbitrário, porque seleciona uma perspectiva única de domínio, excluindo as demais.

Segundo Alvarenga (2009, p.19), há similaridades entre o processo cognitivo de categorização e o processo formal da classificação lógica e, apesar de serem conceitos distintos, se confundem pelo fato de representarem mecanismos utilizados para se estabelecer ordem. A distinção fundamental entre os dois processos reside no modo pelo qual se utiliza o processo na busca do ordenamento. Alvarenga (2009, p.21) ressalta que a classificação tradicional é rigorosa ao ordenar que uma entidade seja ou não membro de uma classe particular. O processo de categorização, no entanto, é flexível e criativo e cria associações não rígidas entre as entidades, baseadas em similaridades observadas dentro de um determinado

contexto. A flexibilidade do processo se caracteriza pelo fato de que o contexto pode variar, alterando, assim, as percepções cognitivas do observador.

A classificação é, portanto, um processo posterior ao da categorização e caracteriza uma tentativa do ser humano de associar os objetos que o cercam a objetos já identificados como pertencentes a um grupo específico. Deste modo, os novos objetos são integrados às estruturas já definidas, para que estas estruturas sejam reorganizadas, facilitando a interação do ser humano com os objetos que o cercam e aperfeiçoando o entendimento das estruturas que compõem o mundo em que vive o observador. Existem vários algoritmos que automaticamente categorizam e classificam os documentos baseados em princípios probabilísticos e geométricos.

2.2 As Patentes

Uma patente é um contrato entre o inventor e o governo, segundo o qual, em troca de divulgação pública e integral de uma invenção, o governo concede ao inventor o direito de excluir terceiros por um período limitado de tempo e de usar ou vender a invenção (HUFKER; ALPERT, 1994, p.45). De acordo com Alberts e outros (2011, p.4), as patentes são documentos jurídicos complexos, que contêm mais detalhes e descrições do que artigos científicos. A linguagem e a formatação do texto da patente são controladas por leis e regulamentos do país ou da autoridade de patentes em que o autor da invenção solicitou a patente.

Patentes contêm resultados valiosos de investigação científica para a indústria, para a economia e para a sociedade. Um documento de patente contém itens estruturados e não estruturados, que podem ser explorados para identificar relações, tendências de negócios e soluções industriais inovadoras (TSENG *et al.*, 2007).

2.3 Inteligência Computacional

A Inteligência Computacional, entendida aqui como um ramo da Inteligência Artificial que se caracteriza pelo uso intensivo de dados, pela relação simbiótica entre o computador e o especialista e pelo emprego de técnicas que apresentam algum tipo de adaptação, aprendizado ou evolução (KORDON, 2010, p.23), é um campo de conhecimento e pesquisa que viabiliza a transferência de processos de representação de conhecimento para o computador, com a proposta de desenvolver sistemas inteligentes capazes de simular o processo de decisão do ser humano (REZENDE, 2003, p.15). Os sistemas inteligentes se caracterizam por apresentar habilidade para usar conhecimento para desempenhar tarefas ou resolver problemas e capacidade para aproveitar associações e inferência para trabalhar com problemas que se

aproximam de situações reais. Os sistemas inteligentes utilizam técnicas que podem ser aplicadas isoladamente ou em conjunto para auxiliar o processo decisório. Dentre as técnicas e metodologias mais utilizadas, podem ser destacadas Aquisição de Conhecimento, Aprendizado de Máquina, Redes Neurais Artificiais, Lógica *Fuzzy*, Computação Evolutiva, Agentes e Multiagentes e Mineração de Dados e de Textos.

As RNA são modelos matemáticos inspirados nas estruturas neurais biológicas e têm capacidade computacional adquirida por meio de aprendizado e generalização (HAYKIN, 1994, p.2). Os trabalhos que utilizam RNA reconhecem que o cérebro trabalha de um modo inteiramente diferente do computador digital convencional. As pessoas armazenam conhecimento por meio de modelos mentais construídos ao longo de suas vidas e adquirem valores éticos e sociais próprios que, dificilmente, poderão ser representados por computadores. Uma RNA é um algoritmo de processamento inspirado no funcionamento das redes biológicas existentes no cérebro, que armazena informação experimental e a torna disponível para uso. Assemelha-se ao cérebro em, basicamente, dois aspectos: o “conhecimento” é adquirido pela RNA por meio de um processo de aprendizagem e as forças de conexão existentes entre os neurônios, conhecidas como pesos sinápticos, são usadas para armazenar este “conhecimento”.

3 METODOLOGIA

O projeto de pesquisa prevê a definição de parâmetros de avaliação para que sejam analisados os sistemas de categorização adotados por diferentes escritórios de patentes e para que seja definido um atributo para a geração de um sistema de categorização de patentes.

Após definido o atributo a ser utilizado no sistema de categorização, propõe-se desenvolver um método de categorização que possibilite a utilização do processo como fonte de informação para o estudo da inovação, da mudança tecnológica e da medida do valor econômico da patente. O trabalho está sendo, basicamente, desenvolvido em duas fases distintas.

A primeira fase é destinada à estruturação das etapas práticas do projeto. Estão sendo realizados os levantamentos dos atributos e das métricas dos processos de categorização de patentes identificados no levantamento bibliográfico e utilizados algoritmos de categorização empregando-se as bases de patentes digitais disponíveis. Nesta fase, estão sendo definidas as métricas e os critérios adequados para avaliar o desempenho de métodos de categorização de patentes.

A segunda fase aprofunda as questões do projeto. Será escolhido um atributo de categorização para ser utilizado e proposto um método de categorização de patentes adequado aos objetivos do projeto. Serão realizados experimentos de validação do método proposto e analisados os resultados e o desempenho do modelo, conforme as métricas definidas.

4 CONSIDERAÇÕES FINAIS

As informações relacionadas às patentes podem ser usadas como uma medida do avanço tecnológico de específicos domínios de conhecimento. Obter conhecimento a partir das informações contidas nas patentes é essencial para processos de tomada de decisão relativos a investimentos de pesquisa. No entanto, patentes são documentos complexos, repletos de detalhes técnicos e exaustivas descrições. Além disso, as diferenças encontradas nos processos relacionados aos escritórios de patentes de diferentes países dificultam a padronização de processos de categorização e de classificação e a definição de métricas para avaliá-los. A proposta deste projeto é criar um sistema de categorização alternativo, que possibilite acesso simplificado às informações incorporadas nas patentes, gerando um formato útil ao processo de recuperação de informação. A informação categorizada torna acessível o conhecimento necessário aos especialistas que podem extrair e analisar o conteúdo comum aos integrantes das categorias.

REFERÊNCIAS

ABADAL, Ernest; CODINA, Lluís. Recuperación de Información. In: _____. **Bases de Datos Documentales: Características, funciones y método.** Madrid: Síntesis, 2005. p. 29-92.

ALBERTS, Doreen; YANG, Cynthia B.; FOBARE-DEPONIO, Denise; KOUBEK, Ken; ROBINS, Suzanne; RODGERS, Matthew; SIMMONS, Edlyn; DEMARCO, Dominic. Introduction to Patent Searching. In: LUPU, M.; TAIT, J.; MAYER, K.; TRIPPE, A. J. **Current Challenges in Patent Information Retrieval.** [S./l.]: Springer, 2011. p. 3-44.

ALVARENGA, Lídia. **Conexões com a “Realidade”: linguagem, classificação e categorização:** notas de aula da disciplina Análises no processo classificatório para Sistemas de Recuperação de Informação, PPGCI – UFMG, 28f, 2009.

ARCHIBUGI, Daniele; PIANTA, Mario. Measuring technological change through patents and innovation. **Technovation**, v. 16, n. 9, p. 451-468, 1996.

CAMUS, Caterina; BRANCALEON, Riccardo. Intellectual assets management: from patents to knowledge. **World Patent Information**, v. 25, n. 2, p. 155-159, 2003.

HAYKIN, Simon. **Neural Networks: a comprehensive foundation.** 2. ed. New Jersey: Prentice Hall, 1994. 842p.

HUFKER, Tim; ALPERT, Frank. Patents: a Managerial Perspective. **Journal of Product and Brand Management**, v. 3, n. 4, p. 44-54, 1994.

JACOB, Elin K. Classification and categorization: a difference that makes a difference. **Library Trends**, v. 52, n. 3, p. 515-540, 2004.

KORDON, Arthur K. **Applying Computational Intelligence**: how to create value. Springer-Verlag, Berlin, 2010. 459p.

LEE, Sungjoo; YOON, Byungun; PARK, Yongtae. An approach to discovering new technology opportunities: Keyword-based patent map approach. **Technovation**, v. 29, p. 481-497. 2009.

LEYDESDORFF, Loet. The University-Industry Knowledge Relationship: Analyzing Patents and the Science Base of Technologies. **Journal of the American Society for Information Science and Technology**, v. 55, n. 11, p. 991-1001, 2004.

MARKELLOS, Konstantinos; PERDIKURI, Katerina; MARKELLOU, Penelope; SIRMAKESSIS, Spiros; MAYRITSAKIS, George; TSAKALIDIS, Athanasios. Knowledge Discovery in Patent Databases. In: XI INTERNATIONAL CONFERENCE ON INFORMATION AND KNOWLEDGE MANAGEMENT (CIKM'02), 11., New York, 2002. **Proceedings...** New York, USA, 2002, p. 672-674.

MEIRELES, Magali Rezende Gouvêa. **Categorização de documentos a partir de suas citações: um método baseado em Redes Neurais Artificiais**. 2012. 237f. Tese (Doutorado em Ciência da Informação) – Escola de Ciência da Informação, Universidade Federal de Minas Gerais. Belo Horizonte. 2012.

REZENDE, Solange Oliveira. **Sistemas Inteligentes**: fundamentos e aplicações. Barueri, SP: Manole, 2003. 525p.

SARACEVIC, Tefko. Ciência da Informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jun. 1996.

TSENG, Yuen-Hsien; LIN, Chi-Jen; LIN, Yu-I. Text mining techniques for patent analysis. **Information Processing and Management**, v. 43, n. 5, p. 1216-1247, 2007.

A FOLKSONOMIA E A REPRESENTAÇÃO COLABORATIVA DA INFORMAÇÃO EM AMBIENTES DIGITAIS

A FOLKSONOMY AND THE COLLABORATIVE INFORMATION REPRESENTATION IN DIGITAL ENVIRONMENTS

Raimunda Fernanda dos Santos
Renato Fernandes Corrêa

Resumo: Discute a Folksonomia e a representação colaborativa da informação em ambiente digital, analisando as práticas colaborativas de representação de conteúdo e suas implicações no processo de recuperação da informação. Investiga a Folksonomia Assistida como uma proposta para maior precisão no processo de busca e recuperação da informação. A metodologia da pesquisa é bibliográfica e exploratória, pautada em revisão de literatura sobre a Folksonomia e a representação colaborativa da informação. Destaca que para minimizar os problemas de representação e recuperação da informação em ambientes colaborativos a alternativa é controlar o nível de liberdade do usuário ao atribuir tag por meio da Folksonomia Assistida. Conclui propondo a aplicabilidade do modelo de representação iterativa no âmbito dos sistemas de bases de dados de periódicos eletrônicos cuja política de indexação permite ao indexador fazer uso dos mesmos termos informados pelos autores nas palavras-chave dos artigos.

Palavras-chave: Folksonomia. Representação colaborativa da informação. Folksonomia Assistida.

Abstract: This work discusses the folksonomy and collaborative representation of information in the digital environment, through analyze of the collaborative practices of representation of content and their implications in the process of information retrieval. It investigates the Assisted Folksonomy as a proposal for improvement of the precision in information retrieval process. The research methodology is exploratory and bibliographical, and consists of literature revision about folksonomy and collaborative information representation. It highlights that the alternative to minimize the problems of representation and retrieval of information in collaborative environments is to control the level of user freedom in assignment of tag by Assisted Folksonomy. Concludes by proposing the applicability of the iterative representation model within the electronic journal database systems whose political of indexing allow to the indexer use the same keywords informed by the authors of the articles.

Keywords: Folksonomy. Collaborative information representation. Assisted Folksonomy.

1 INTRODUÇÃO

No hodierno contexto informacional, a Folksonomia é utilizada na categorização de conteúdos disponíveis na rede e nos repositórios digitais, consolidando, assim, seu papel na indexação colaborativa (etiquetagem) de recursos informacionais. Nesse sentido, as práticas colaborativas de organização da informação na Web são assim incentivadas fazendo com que as pessoas contribuam para a formação de uma inteligência coletiva, uma tendência de representar a informação através da livre expressão e do entendimento dos usuários da Web a respeito de determinado conteúdo.

No âmbito da Ciência da Informação o tema Folksonomia ainda é pouco explorado, principalmente no que concerne à sua aplicação em sistemas de informação, onde os seus usos e implicações propõem inúmeras possibilidades de discussões. Nesse contexto, e visando a participação direta do usuário na organização e representação de conteúdos em ambientes digitais colaborativos, se constitui como objetivo do presente trabalho investigar a Folksonomia e a representação colaborativa da informação em ambiente digital analisando as práticas colaborativas de representação de conteúdo e as suas implicações no processo de recuperação da informação, bem como discutir a Folksonomia Assistida como uma proposta para melhorar a estrutura de representação do conhecimento no âmbito dos sistemas colaborativos de informação.

Quanto aos objetivos a serem alcançados para este estudo, foi empreendida uma pesquisa do tipo exploratória. Quanto aos meios, os procedimentos metodológicos foram empreendidos a partir de uma pesquisa bibliográfica, baseada em revisão de literatura sobre Folksonomia e a representação colaborativa da informação. Foram utilizadas na concepção do trabalho fontes impressas como livros, e em meio digital como artigos de periódicos científicos eletrônicos e dissertações da área de Ciência da Informação encontradas por meio de bases de dados, mecanismos de busca e bibliotecas digitais na Web. Mediante tais considerações, acredita-se que a presente pesquisa é relevante uma vez que irá contribuir com a produção científica sobre Folksonomia na literatura nacional da Ciência da Informação, bem como trazer à discussão sua aplicabilidade no contexto dos sistemas colaborativos de informação.

2 A FOLKSONOMIA E AS SUAS IMPLICAÇÕES NO PROCESSO DE RECUPERAÇÃO DA INFORMAÇÃO

No contexto da Web social, a Folksonomia está inserida diante da nova interface entre usuários e informação em rede. De acordo com Amstel (2007) ela faz parte de uma terminologia cunhada por Thomas Vander Wal numa lista de discussão sobre arquitetura da informação em 2004. Na lista de discussão, Eric Sheid propôs o termo “*folkclassification*” e Thomas Wander Wall complementou com “*folksonomy*”, formando o termo “*folk*” do germânico “povo” e “*taxonomy*”, do grego “regra de revisão”. Assim, a Folksonomia é uma analogia ao termo taxonomia¹⁹⁰ tendo como principal característica a criação de *tags* com função de descritor a partir da linguagem dos usuários. De acordo com Blattmann e Silva (2007) a Folksonomia é uma “[...] forma relacional de categorizar e classificar informações

¹⁹⁰ Estudo da classificação das coisas. É o ato de dar nomes, classificar, identificar.

disponíveis na *Web*, sejam elas representadas por meio de textos, imagens, áudio, vídeo ou qualquer outro formato”. Ao invés de utilizar uma forma hierárquica e centralizada de categorização de alguma coisa, o usuário escolhe palavras-chave (conhecidas como “*tags*”¹⁹¹) para classificar a informação.

Destarte, levando em conta a literatura brasileira em ciência da informação sobre Folksonomia, percebe-se que são discutidos na maioria dos trabalhos aspectos relativos à ausência de uma linguagem de indexação controlada para representar os objetos informacionais nas ferramentas da denominada Web 2.0 (Web social). Tais ideias partem do princípio de que o alto grau de liberdade para a categorização pode acentuar a descentralização no processo, pois quem classifica o conteúdo são as próprias pessoas interessadas neste, com diversos graus de envolvimento tanto de conhecimento, quanto por experiência ou sentimento, constituindo assim uma indexação social - o que gera implicações no processo de recuperação da informação em sistemas de informação. À guisa de exemplos, é possível perceber a presença de vários sítios da Web 2.0 que usam Folksonomia, tais como o *Flickr* (álbum pessoal de fotos), o *You Tube* (vídeos), *Instagram* (rede social de compartilhamento de fotos), *Facebook* (rede social de relacionamentos), dentre outros.

Noruzi (2007) em sua pesquisa menciona algumas implicações da Folksonomia no processo de representação da informação em ambientes colaborativos, tais como a ocorrência de plurais¹⁹², existência de polissemia¹⁹³, emprego de sinonímia¹⁹⁴, profundidade/especificidade da descrição¹⁹⁵.

Além dos problemas supracitados, Amstel (2007), Gouvêa; Loh (2007) e Catarino; Baptista (2009); enfatizam que em ambientes que utilizam a representação colaborativa da informação é comum a presença de descritores **homógrafos**¹⁹⁶; **erros ortográficos**; **expressões regionais**¹⁹⁷; **etiquetas em diferentes idiomas e com letras maiúsculas e**

¹⁹¹ Em português significa “etiqueta”, “identificação”. “Taggear” é identificar, etiquetar alguma coisa.

¹⁹² Ocorrência de palavras com flexões gramaticais de números e variações lingüísticas de gêneros nos descritores utilizados.

¹⁹³ Palavra que possui mais de um significado.

¹⁹⁴ Palavras distintas que possuem o mesmo significado.

¹⁹⁵ Refere-se à especificidade dos descritores elencados pelo usuário no processo de representar o conteúdo do objeto informacional em tags. Tais descritores (em sua maioria) podem não representar o objeto em sua totalidade, sendo elencados apenas para interesse particular do usuário que o atribuiu.

¹⁹⁶ Termos que coincidem na sua forma, porém representam conceitos diferentes.

¹⁹⁷ Expressões empregadas como *tags* de acordo com o contexto social/regional em que o usuário estar inserido.

minúsculas (sem padronização na descrição) e **etiquetas assimétricas**¹⁹⁸. Levando em consideração os aspectos teóricos enfatizados anteriormente e objetivando revisar tais questões na prática, será apresentada na Figura 2 uma imagem disponível no *Facebook*¹⁹⁹ com as suas respectivas tags empregadas pelo usuário do sistema.

FIGURA 2 - Atribuição de tags em foto disponível no facebook



Fonte: <https://www.facebook.com/>

Como pode ser verificado, o usuário da rede social supracitada elencou 23 descritores para esta imagem, são eles: "goodnight", "anoitecer", "crepusculo", "interior", "florania", "cielo", "sky", "ceu", "nuages", "lua", "sun", "sol", "coqueiros", "curtofloraniarn", "forrogram", "nordestegram", "brazil", "cidadedasflores", "love", "amomuito", "estrela", "mtut", "brasilgram". Diante disso, percebe-se que o usuário que postou este objeto informacional o representou por meio de conceitos que foram ativados em sua estrutura mental ou identidade cultural em relação à imagem representada, destacando até mesmo seus gostos pessoais - como verificado nas tags "amomuito" "curtoflorania" "love". Percebe-se, nesse sentido, que a forma de escrita dos descritores é subjetiva fazendo sentido (em sua maioria) apenas para quem os atribuiu. Sob esse entendimento percebe-se que, diante desta flexibilidade na representação da informação, são agregadas as manifestações da linguagem natural²⁰⁰ em ambiente digital tendo em vista que o usuário, ao rotular conteúdos, expressa a

¹⁹⁸ Tags que não possuem relações entre si.

¹⁹⁹ Rede social lançada em 04 de fevereiro de 2004, constitui em um serviço em que os usuários compartilham conteúdos e interagem entre si.

²⁰⁰ Consiste em um discurso comum. Uma linguagem utilizada habitualmente na escrita e na fala. Diferentemente das linguagens documentárias ou vocabulários controlados que consistem em linguagens construídas de acordo com a necessidade dos usuários (LANCASTER, 2004).

sua estrutura mental ou identidade cultural em relação ao objeto a ser representado. Portanto, é importante enfatizar que a ausência de controle desses fenômenos pode diminuir a revocação e precisão²⁰¹ durante o processo de recuperação da informação.

Neste contexto, surge na literatura a proposta de Folksonomia Assistida com o objetivo de buscar melhorar a eficiência do uso de *tags* para a representação da informação, permitindo ao usuário uma descrição livre para os objetos digitais, porém orientada pelo próprio conhecimento disponibilizado no sistema de informação utilizado, conforme será discutido na próxima seção.

3 FOLKSONOMIA ASSISTIDA: UMA NOVA POSSIBILIDADE DE REPRESENTAÇÃO DA INFORMAÇÃO EM AMBIENTE DIGITAL

Tendo em vista a participação do usuário no processo de etiquetagem dos objetos informacionais em repositórios digitais científicos, Santarém Segundo (2010b) destaca a Folksonomia como elemento primordial para a construção de um modelo de representação iterativa cujo processo ele denomina de Folksonomia Assistida²⁰². Para este autor, o modelo de representação iterativa parte do princípio da existência de uma *interface* diferente no momento do usuário informar as palavras-chave que serão utilizadas como descritores que também passarão a ficar disponíveis como sugestões de *tags* para serem empregadas por outros usuários. Nesse entendimento, a iteratividade se faz presente no momento da retroalimentação da informação por meio do enriquecimento e fortalecimento do conjunto de descritores visando a reconstrução da estrutura de representação do conhecimento do repositório digital de forma contínua. Assim, a Folksonomia Assistida objetiva qualificar e melhorar a estrutura de representação do conhecimento em ambientes colaborativos, possibilitando que *as tags* tenham um maior significado em relação ao objeto informacional tratado, principalmente no contexto em que o mesmo está inserido "de forma que o usuário do sistema evite abreviações, plurais/singulares ou ainda palavras que possam dificultar a recuperação da informação, posteriormente" (SANTARÉM SEGUNDO, 2010a, p.181).

Inicialmente no momento da descrição das *tags*, o sistema deve receber os termos elencados pelo usuário para realizar a pesquisa de relacionamento da informação em estruturas de representação do conhecimento. Neste momento, o sistema deve apresentar um conjunto de informações já inseridas pelo usuário administrador (com o auxílio de

²⁰¹ De acordo com Lancaster (2004), estas medidas de desempenho visam medir a capacidade de recuperar documentos úteis e evitar documentos inúteis, baseando-se na noção de relevância.

²⁰² Do inglês Assisted Foksonomy.

vocabulários controlados) como uma sugestão de *tags* que poderão ser aceitas de forma total ou parcial. Portanto, ressalta-se que a aplicação de termos inerentes às estruturas de representação do conhecimento (taxonomias, tesouros e ontologias) não possibilita “impedir” ou “engessar” a criatividade do usuário, nem descaracterizar o termo Folksonomia pois, conforme citado anteriormente, o sistema deve possibilitar que o usuário decida, de forma autônoma, os termos que poderão ser utilizados como descritores de forma que ele possa usar a base informacional deste ambiente para qualificar a descrição de seu recurso, ou seja, é uma espécie de “Folksonomia Controlada” que auxilia o usuário na descrição da *tag* assunto do seu objeto informacional.

Esta possibilidade de representação da informação é realizada no âmbito do Faviki²⁰³ que consiste em um espaço social de compartilhamento de links ou marcadores favoritos (*bookmarks*), e que possibilita a utilização de conceitos da Wikipedia como *tags*. Nesse sentido, quando o usuário atribui uma etiqueta neste ambiente, automaticamente ele recebe sugestões de *tags* relacionadas, nomeadas de "etiquetas semânticas". Através da utilização deste modelo, o Faviki livra-se dos problemas criados pelas denominadas **etiquetas assimétricas** (ou seja, os descritores que não possuem relações entre si), conforme ilustrado na FIGURA 3.

FIGURA 3 - Atribuição de tags no Faviki



Fonte: <http://www.faviki.com/pages/welcome/>

Na FIGURA 3, após o usuário do Faviki ter digitado a tag "webdesign" para uma página da Web, é apresentado de forma sistemática conceitos da Wikipedia previamente

²⁰³ Disponível em: <<http://www.faviki.com/pages/welcome/>>. Acesso em: 13 jul. 2014.

disponíveis para que o mesmo possa verificar se o descritor possui relação semântica com o que será representado. O Faviki também sugere como *tags* os conceitos da Wikipedia encontrados no corpo da página web que está sendo adicionada à base. Mesmo que a *tag* digitada seja sintaticamente diferente do texto especificado como conceito, o sistema permite associar a *tag* ao conceito correto via URL.

Mediante tais considerações, para se alcançar precisão no processo de representação e recuperação da informação em sistemas de informação colaborativos, Reis (2007) afirma que a Folksonomia Assistida pode ser implementada de várias formas, dentre elas: apresentar ao usuário instruções para escolha das *tags*; sugerir descritores semelhantes já utilizados por outros usuários; apresentar sinônimos, aplicar corretores ortográficos, indicar o assunto ao qual a *tag* se refere e sugerir ao usuário que classifique a informação em descritores pré-definidos. Além de tais aspectos, pode-se inferir também a necessidade do empenho do profissional da informação com o auxílio de profissionais de outras áreas do conhecimento, para que sejam desenvolvidos instrumentos de controle de vocabulário e procedimentos que permitam diminuir alguns problemas relativos à não normalização das *tags* em ambientes colaborativos.

4 CONSIDERAÇÕES FINAIS

Diante do exposto, é possível perceber a importância do uso da Folksonomia para a classificação e a organização de conteúdos disponíveis em ambientes digitais. Contudo, para minimizar os problemas de representação e recuperação da informação em ambientes que utilizam a indexação colaborativa, a alternativa é controlar o nível de liberdade do usuário ao atribuir *tags* através da Folksonomia Assistida. Dessa forma, o processo de iteratividade resulta na reconstrução do conhecimento, de forma coletiva e “moderada”, permitindo uma nova estrutura de organização e representação de informação mais consistente para o domínio em que o sistema de informação está inserido.

Levando em consideração que a Folksonomia também está presente no âmbito das bases de dados de artigos de periódicos eletrônicos, acredita-se que o modelo de representação iterativa por meio da Folksonomia Assistida, pode ser aperfeiçoado e aplicado no âmbito dos sistemas de bases de dados de periódicos eletrônicos cuja política de indexação permite ao indexador atribuir para a produção intelectual depositada os termos informados pelos autores nas palavras-chave dos resumos de seus artigos. Mediante tais considerações, como trabalhos futuros será analisado de forma mais detalhada o uso da Folksonomia Assistida em sites da Web e, em seguida, será proposta uma metodologia para implantação da

Folksonomia Assistida na Base de Dados Referencial de Artigos de Periódicos da área de Ciência da Informação (BRAPCI) objetivando melhorar a representação, a recuperação e o acesso à informação contida neste ambiente. Por fim, é importante ressaltar que a escolha desta fonte de informação deve-se à importância que representa para a Ciência da Informação no Brasil e a sua seleção está atrelada à sua política de indexação.

REFERÊNCIAS

- AMSTEL, Van Frederick. **Folksonomia**: vocabulário descontrolado na arquitetura da informação ou samba do crioulo doido. 2007. Disponível em: <http://www.guilhermo.com/ai_biblioteca/referencia.asp?referencia=317> . Acesso em: 20 maio 2014.
- BLATTMANN, Úrsula ; SILVA, Fabiano Couto Corrêa da. Colaboração e interação na Web 2.0 e biblioteca 2.0. **Revista ACB** : biblioteconomia em Santa Catarina, v.12, n.2, jul./dez. 2007. Disponível em: <<http://www.asbsc.org.br/revista/ojs/viewissue.php?id=21#Artigos>>. Acesso em: 18 jun. 2014.
- CATARINO, Maria Elisabete; BAPTISTA, Ana Alice. Folksonomias: características da etiquetas na descrição de recursos da Web. **Informação & Informação**, Londrina, v. 14, n. esp., 2009.
- GOUVÊA, Cleber; LOH, Stanley. Folksonomias: identificações de padrões na seleção de tags para descrever conteúdos. **Revista Eletrônica de Sistemas de Informação**, vol. 11, n.2, 2007. Disponível em: <<file:///C:/Users/fernanda/Dropbox/MESTRADO/DISSERTA%C3%87%C3%83O/Sobre%20Folksonomia/Gouv%C3%AAs%20e%20Loh.pdf>>. Acesso em: 24 jul. 2014.
- LANCASTER, F. W. **Indexação e resumos**: teoria e prática. Brasília: Briquet de Lemos, 2004.
- NORUZI, A. Folksonomies: Why do we need controlled vocabulary? **Webology**, v.4, n.2, 2007. Disponível em: < <http://www.webology.org/2007/v4n2/editorial12.html>>. Acesso em: 20 ago. 2013.
- REIS, Guilherme. **Arquitetura da Informação de websites**. 2007. Disponível em: < <http://www.guilhermo.com/apresentacoes/arquiteturainformacao-fesa-guilhermoreis.pdf>>. Acesso em: 24 jul. 2014.
- SANTARÉM SEGUNDO, José Eduardo. Rede de tags para recuperação da informação no contexto da Representação Iterativa. **InCID: Rev. Ci. Inf. e Doc.**, Ribeirão Preto , v.2, n.1, p.86-109, jan./jun. 2010a.
- SANTARÉM SEGUNDO, José Eduardo. **Representação Iterativa**: um modelo para repositórios digitais. Marília,SP, 2010. 224f. Tese (Doutorado em Ciência da Informação) – Universidade Estadual Paulista Júlio de Mesquita Filho, Marília, SP, 2010b.

TESAURO DO REGIME MILITAR BRASILEIRO
BRAZILIAN MILITARY GOVERNMENT YEARS' THESAURUS

Renato Rocha Souza
 Suemi Higuchi
 Alexandre Morelli
 Angela Moreira
 Bernardo de Paola Bortolotti Faria
 Bianca de Magalhães Silveira
 Celso Castro
 Daniele Chaves Amado
 Juliana Gagliardi
 Letícia Carvalho Ferreira
 Marco Aurélio Vannuchi Leme de Mattos
 Martina Spohr
 Renan Marinho de Castro
 Vanessa Matheus Cavalcante

Resumo: O presente artigo apresenta projeto conjunto do CPDOC (Centro de Pesquisa e Documentação de História Contemporânea do Brasil) e da Escola de Matemática Aplicada, ambos da Fundação Getúlio Vargas, com o objetivo de construção de tesouros retratando período da história contemporânea no Brasil. Como primeiro produto, apresenta-se tesouro que retrata o período do Regime Militar Brasileiro – também conhecido como período de ditadura militar – construído à partir dos verbetes do Dicionário Histórico Biográfico Brasileiro; de material existente nos Arquivos Pessoais (conjunto documental sobre personalidades da história contemporânea brasileira) e das entrevistas em áudio e vídeo pertencentes ao Programa de História Oral, todos estes fazendo parte do acervo documental do CPDOC. No artigo são apresentadas as características do acervo utilizado, a motivação da construção, as categorias básicas do tesouro, assim como a equipe, tecnologias empregadas e o processo de construção. Os resultados alcançados até o momento são apresentados, assim como as direções futuras de pesquisa, que apontam para o desenvolvimento de thesauri que cubram outros períodos da história contemporânea no Brasil.

Palavras-chave: Regime Militar Brasileiro, Ditadura Militar, Tesouro.

Abstract: This paper presents a joint project between the CPDOC (Center for Teaching and Research in the Social Sciences and Contemporary History of Brazil) and the Applied Mathematics School, both from the Getúlio Vargas Foundation. This project aims at building thesauri in the contemporary Brazilian history domain. As a first product, we are presenting a thesaurus on the Brazilian Military Government, also known as “military dictatorship period”. The documental sources for building the thesauri are parts of the CPDOC archives, as the “Personal Archives”, that comprises documents on personalities of the Brazilian contemporary history; the “Brazilian Historical Biographical Dictionary” and the “Oral History Program”. In this paper we present the characteristics of the documents and sources, the motivation for building such work, and the main thesaurus categories and facets, along with the multidisciplinary team, technologies and the building process. The results obtained so far and the future research directions are also presented, mainly on the development of thesauri covering other periods of the Brazilian recent history.

Keywords: Brazilian Military Government, Military Dictatorship, Thesauri.

1 INTRODUÇÃO

O Centro de Pesquisa e Documentação de História Contemporânea do Brasil (CPDOC) da Fundação Getúlio Vargas (FGV) é considerado um centro de referência dedicado ao estudo e à preservação da memória do país – particularmente dos anos 1930 em diante -, e é conhecido por abrigar e disponibilizar um importante acervo histórico documental. Este acervo é constituído de diferentes tipos de informação distribuídos em sistemas próprios, construídos de acordo com as especificações e necessidades demandadas por cada um destes tipos. Por se tratar de um mesmo universo de discurso, a saber, a história recente do Brasil, em um sistema heterogêneo e distribuído de arquivos, o uso de instrumentos de controle terminológico faz-se necessário, pois é o que permitirá a seleção do conjunto de conceitos relevantes capazes de traduzir adequadamente o conteúdo das unidades documentais armazenadas nestes sistemas, de forma integrada e padronizada. Este artigo apresenta de maneira geral o projeto de criação de Tesouros Terminológicos que o CPDOC vem desenvolvendo, visando a aprimorar o tratamento das informações e, conseqüentemente, a estratégia de busca e interação em seu sistema de recuperação da informação. Especificamente, detalha o primeiro produto deste projeto, que é o Tesouro do Regime Militar Brasileiro.

2 O CPDOC E SEU ACERVO

O Acervo do CPDOC se divide de maneira geral em três sistemas: a) de Arquivos Pessoais; b) do Programa de História Oral (PHO) e c) do Dicionário Histórico-Biográfico Brasileiro (DHBB). O primeiro reúne informações de um importante conjunto de arquivos pessoais, considerado como dos mais expressivos no país. São documentos de valor inestimável na forma de manuscritos, diários, cartas, discursos, filmes e fotografias, acumulados por indivíduos que atuaram em funções diversas no cenário nacional. Dentre estes, encontram-se presidentes, ministros, militares, líderes e diplomatas. Atualmente o acervo é constituído de aproximadamente 200 fundos que somam cerca de 1,8 milhões de documentos. Já o Programa de História Oral (PHO) abriga um importante e diversificado banco de depoimentos. Criado em 1975 a partir de metodologia pioneira em âmbito nacional, o programa produziu neste período mais de seis mil horas de gravação correspondentes a cerca de duas mil entrevistas. Inicialmente, tinha como objetivo principal complementar a pesquisa documental através de entrevistas com personalidades políticas ligadas direta ou indiretamente ao Estado brasileiro (políticos, intelectuais, tecnocratas, militares e diplomatas, entre outros). Mas hoje contempla também temáticas ligadas à história da imprensa,

movimentos sociais, elite empresarial e história do esporte. O terceiro dos sistemas mantém informações do Dicionário Histórico-Biográfico Brasileiro (DHBB), obra de referência voltada para a história do Brasil pós-1930 que, na versão atual, abrange mais de 7.500 verbetes de natureza biográfica e temática. São pessoas, instituições, organizações e eventos cuidadosamente selecionados a partir de critérios que mensuram a relevância destes para a história política do período, e que constituem pontos de acesso a textos que relatam a história, origem, definições e demais esclarecimentos do conceito que representam. A obra encontra-se em permanente processo de atualização e pode, juntamente com os outros acervos, ser consultada na íntegra a partir do portal do CPDOC²⁰⁴.

3 O PROBLEMA

O tratamento da informação no CPDOC adota o uso de vocabulários controlados para descrever o seu universo temático e realizar o controle das variações terminológicas nos descritores dos sistemas. Estes são os principais instrumentos de indexação utilizados para organização dos acervos de Arquivos Pessoais e de História Oral. Apesar da automação dos sistemas e da existência de múltiplos pontos de acesso aos itens documentais, essa forma de indexação ainda é bastante limitada e tem causado dificuldades nos níveis de gestão e de recuperação da informação. Por adotarem a estrutura de glossários, não estabelecem nenhuma hierarquia, categorização ou qualquer outro tipo de relacionamento entre os termos, salvo a distinção entre descritores “eleitos” e “não eleitos” - dentre os termos equivalentes. Em última análise, a ausência de conexões entre conceitos não permite navegação hierárquica nem a busca por associação. Outro ponto importante a ser considerado diz respeito à utilização de diferentes vocabulários controlados para referenciar os mesmos assuntos nos três sistemas, devido à adoção de critérios idiossincráticos. Este fato é agravado pela inexistência de notas de aplicação ou de qualificadores que delimitem o significado dos termos, além de gestão independente em cada um dos sistemas. Para os usuários e pesquisadores em busca de informações, esse problema terminológico pode diminuir as taxas de revocação e precisão nas buscas transversais entre os acervos. Em suma, os problemas são os seguintes:

- Existência de vocabulários conflitantes
- Ausência de hierarquias ou associações
- Não há exaustividade nos instrumentos terminológicos para descrever todo o conjunto de conceitos presente no domínio.

204

Disponível em <http://www.fgv.br/cpdoc/busca>

4 A SOLUÇÃO PROPOSTA

Como uma possível solução, foi proposta a criação de tesouros temáticos, de forma a cobrir os períodos da história brasileira recente. A criação dos tesouros possibilitaria o enriquecimento dos vocabulários através da explicitação das relações conceituais entre os termos e permitiria a criação de categorias primárias para classificar os tipos de conceitos, como personalidades, eventos e locais. A divisão por períodos históricos se justifica pelo alto grau de adesão, entre os pesquisadores, aos recortes temporais que estabelecem os períodos a analisar. Neste sentido, escolheu-se como tema piloto o período do regime militar brasileiro, conhecido como “ditadura militar”, que compreende os anos de 1964 a 1985. Esta escolha se justifica pela relevância histórica dos registros pertencentes ao acervo que datam deste período e a visibilidade para a sociedade do produto gerado. Além disso, a ubiquidade deste tema nos sistemas de informação do CPDOC facilita o processo de sistematização. Em um prévio levantamento realizado na Web, verificou-se a inexistência de tesouros na área da história do Brasil - em especial do período contemporâneo - que pudessem apoiar o projeto, o que motivou ainda mais a equipe no prosseguimento dos trabalhos.

No processo de decisão sobre as formas de representação do conhecimento adequadas, chegou-se a cogitar o desenvolvimento de ontologias. Os tesouros foram escolhidos por apresentarem um ciclo de desenvolvimento mais célere, além de proporcionarem suporte adequado ao processo de indexação. Tesouros apresentam formas alfabéticas e sistemáticas de exibição, e contêm uma estrutura lógica que contempla as relações hierárquicas (do conceito geral para o específico), de equivalência (sinônimos e quase sinônimos) e associativas (relacionamentos transversais) entre os termos, resultando em informação mais precisa e com menos ambiguidade (GOMES, 1990; CAMPOS, 2002).

5 METODOLOGIA

Primeiramente, foi criado um grupo composto por profissionais do CPDOC e da Escola de Matemática Aplicada da FGV. Dentre as formações, historiadores, cientistas sociais, arquivistas, bibliotecários e cientistas da informação, caracterizando um ambiente de trabalho interdisciplinar, tal como a tarefa demanda. Adotou-se para construção do tesouro a norma ANSI/NISO Z39.19-2005, elaborada pela *American National Standards Institute/ National Information Standards Organization* (ANSI/NISO, 2005). Esta define tesouros como vocabulários controlados dispostos em ordens específicas e estruturados de forma que os vários relacionamentos entre os termos - de equivalência, hierárquicos e associativos - sejam exibidos de forma clara e identificados por indicadores padronizados de relacionamentos.

A metodologia preconizada pela norma indica os seguintes passos:

- 1 Pesquisar os vocabulários existentes, de forma a evitar a duplicação e o retrabalho
- 2 Determinar a estrutura e os formatos de exibição
- 3 Escolher o método de construção, dentre as três possibilidades: Abordagem de Comitê, Abordagem Empírica e Abordagem Indutiva
- 4 Registro dos termos e seus relacionamentos (lógicos - hierárquicos e ontológicos - partitivos e associativos).
- 5 Validação dos termos
- 6 Escolha dos termos candidatos (a fazer parte do produto final)
- 7 Escolha dos níveis de especificidade
- 8 Escolha de termos não associados (para garantir a estrutura do tesouro)

Como primeira etapa, delimitou-se o escopo temático do tesouro e os critérios de inclusão dos termos. Esta delimitação levou em consideração a organização atual das informações nos sistemas do CPDOC, o que ensejou a estratégia de elaboração de bases terminológicas integradas.

O formato escolhido de tesouro é o conceitual, com formas sistemáticas e alfabéticas. Estes tipos de tesouros são baseados nas teorias do conceito, da terminologia e da classificação, que fornecem os princípios fundamentais para a construção do vocabulário. O método de categorização ou facetagem, desenvolvido pelo matemático e bibliotecário indiano Shyali Ranganathan na década de 1920, permite agrupar os conceitos de mesma natureza em classes gerais ou facetas, sendo adotado para o estabelecimento das bases para a seleção dos termos do domínio.

A abordagem escolhida para a construção foi híbrida, com forte viés da abordagem de comitê. A seleção dos termos se deu primariamente de forma “bottom up”, em que se recorreu principalmente às duas principais fontes conceituais disponíveis: a lista de termos controlados utilizada nos sistemas de Arquivos Pessoais e História Oral, compreendendo cerca de 12 mil descritores, e as entradas dos quase 7.500 verbetes do Dicionário Histórico-Biográfico Brasileiro²⁰⁵. Em seguida, através do exame dos tipos de conceitos contidos nestes vocabulários, foram propostas as categorias, a seguir:

205

<http://cpdoc.fgv.br/acervo/dhbb>

- Personalidades

A categoria abrange pessoas que tiveram alguma participação direta ou indireta relativa ao Regime Militar Brasileiro, seja através de posicionamento político, ideológico ou ativo no interior de instituições ou na sociedade civil. A equipe considerou que a categoria Personalidades não contemplaria a criação de subcategorias uma vez que a associação dos seus indivíduos com indivíduos da classe Papéis Desempenhados seria recurso suficiente para designar o tipo de personalidade no domínio em questão.

- Instituições/ Organizações

Instituições são estruturas parciais da sociedade que desempenham funções específicas na vida social. Possuem certas normas que tendem a ser reconhecidas ou impostas socialmente (Birou, 1982). Consideram-se nesta categoria as instituições que tenham tido alguma participação direta ou indireta com o Regime Militar Brasileiro, sejam estas políticas, governamentais ou não, da sociedade civil organizada, militares, paramilitares, e ainda aquelas perseguidas, fechadas, ameaçadas, intimidadas, que tenham se oposto ou apoiado ao regime, declaradamente ou não. Já organizações são entidades geradas a partir da combinação de esforços individuais, que têm por finalidade realizar propósitos coletivos de interesse social, político, econômico, mas que não possuem o status de instituições. Da mesma forma que essas, consideramos para fins de inclusão no Tesouro as organizações que se relacionam de alguma forma com o regime militar, sendo parte formal da estrutura do governo, fonte de apoio a este, ainda que de forma independente, ou fonte de oposição ao mesmo.

Foram estabelecidas as seguintes classes para esta categoria: a) Instituições/Organizações estatais e b) Instituições/Organizações da sociedade civil.

- Documentos

A categoria documentos compreende todo registro textual, sonoro, audiovisual ou iconográfico que guarde relações com os processos, eventos ou papéis pertinentes ao domínio, desde que nomeado e de domínio público. Nesta categoria podem-se incluir, por exemplo, atos institucionais, discursos, canções, etc.

- Papéis desempenhados

Compreende todo o rol de funções, oficiais ou não, desempenhadas por indivíduos no contexto do Regime Militar. Entre as atribuições oficiais, consideramos aquelas relacionadas à ocupação de cargos nos poderes Executivo (presidente, ministros), Legislativo (senadores e deputados federais) e Judiciário (ministros), e nas Forças Armadas (Aeronáutica, Exército, Marinha). As funções não oficiais envolvem participantes de movimentos políticos e sociais,

profissionais liberais (empresários, jornalistas) e membros da Igreja favoráveis ou não ao Regime, desde que tenham reconhecidamente assumido parte no período do Regime Militar Brasileiro

- Eventos/Processos

Eventos compreendem acontecimentos ou ocasiões extraordinárias com uma temática particular que gere desdobramentos específicos previstos ou não, e relacionados de alguma forma com a existência do regime militar para lhe dar sustentação ou lhe fazer oposição. Como exemplos, podemos citar a Marcha da Família com Deus pela Liberdade, a Guerrilha do Araguaia, a Passeata dos cem mil e a cassação de mandatos específicos. Considera-se processo como a realização de alguma atividade ou ação de forma contínua e prolongada no tempo, sem que haja parâmetros definitivos para definir datas de início e término. Estão incluídas nesta categoria, além de métodos práticos e diretos adotados em sequência - como a censura, repressão, tortura e a espionagem - linhas de pensamento ou ideologias que também implicam, de uma forma mais geral, em atividade contínuas ao longo do tempo, como o socialismo, o nacionalismo, comunismo e o anticomunismo.

- Locais

Adotando uma perspectiva derivada das categorias fundamentais de Ranganathan (1967), “locais” equivale à categoria “Space” do PMEST, ou seja, abrange unidades geográficas como bairros, cidades, estados e países; superfícies, espaços interiores e exteriores (ex: sala, jardim, escritório); características fisiográficas em geral que, no caso específico, guarde relações com eventos relevantes no escopo do tesouro. Como exemplos destes locais podemos citar: Cinelândia, Ilha das Cobras, Riocentro.

- Períodos

Como colocado, e para fins de maximização da informação provida, independentemente de uma posição específica quanto à duração do período do Regime Militar Brasileiro, o período englobado no tesouro foi fixado como tendo início na renúncia de Jânio Quadros, em 1961, e fim na promulgação da Constituição, em 1988. Quaisquer partições temporais e relevantes deste período serão consideradas para inclusão no tesouro.

Esta categorização foi fortemente influenciada pelo modelo PMEST de Ranganathan (1967). Acredita-se, porém, que, no decorrer do processo de confecção do Tesouro, este primeiro esquema de categorias possa ser modificado de modo a melhor se adequar ao conjunto de conceitos. Os termos vêm sendo analisados sequencialmente a partir dos vocabulários de origem, sendo que a grande maioria se compunha de personagens históricos.

O grupo de analistas têm-se dividido em grupos de trabalho a partir de suas especialidades e sistemas de informação em que trabalham (Dicionário Histórico Biográfico Brasileiro, Arquivos Pessoais, Historia Oral), e o cotejamento das estruturas propostas se dá em reuniões esporádicas, quando são aprovadas as inserções e debatidas as escolhas terminológicas e conceituais. A partir daí, efetua-se o registro dos termos e relacionamentos aprovados.

Para registro do tesouro, estudaram-se algumas ferramentas, como o TCS10²⁰⁶, que havia sido utilizado em trabalhos anteriores, mas descobriu-se que o produto foi descontinuado. Cogitou-se utilizar o software Protégé²⁰⁷, utilizado para construção de ontologias, mas as ferramentas de suporte ao padrão SKOS²⁰⁸, preconizado pelo W3C para publicação de tesouros e outros sistemas de organização do conhecimento no escopo da Web Semântica, ainda não se mostraram satisfatórias. Após uma exaustiva pesquisa e teste de muitas das opções disponíveis²⁰⁹, decidiu-se pelo uso do software Tematres²¹⁰ que, além de gratuito, é *open source* e oferece uma grande gama de opções de exportação e visualização.

6 RESULTADOS INICIAIS E TRABALHOS FUTUROS

No presente momento há categorias bastante populadas, e conta-se com cerca de 800 termos alimentados com suas notas de escopo. Espera-se que o tesouro contenha alguns milhares de termos na sua versão final, tendo sido aplicados de forma adequada os relacionamentos entre eles. Como subprodutos, espera-se a publicação dos tesouros temáticos na página do CPDOC²¹¹ tanto estruturado como um produto independente, quanto na forma de uma taxonomia de pontos de acesso para os sistemas do CPDOC. Desta forma, será proposta a unificação das listas de descritores, que serão, a partir de então, utilizados para a indexação e preenchimento de campos de metadados temáticos. Como uma etapa posterior, estes tesouros servirão de base conceitual e relacional para a criação de ontologias, que permitirão a unificação dos sistemas do CPDOC em um portal semântico, com metadados na forma de uma triplestore.

7 CONSIDERAÇÕES FINAIS

206 <http://www.greymatter.com/product/WebChoir-Inc/Webchoir-TCS-10/278960>

207 <http://protege.stanford.edu/>

208 <http://www.w3.org/2004/02/skos/>

209 <http://www.willpowerinfo.co.uk/thessoft.htm>

210 <http://www.r020.com.ar/tematres/>

211 <http://cpdoc.fgv.br/>

Ao prepararmos um tesouro sobre o Regime Militar Brasileiro tivemos que tomar uma série de decisões que precisam ser explicitadas. Em primeiro lugar, assumimos a centralidade, para o projeto, do acervo documental e informacional constituído pelo CPDOC ao longo de 40 anos, embora sem se limitar a ele. Desse modo, tomamos como ponto de partida o esforço de cotejar e conciliar três diferentes "linguagens controladas" que já adotávamos: a que servia para a indexação de documentos arquivísticos, a utilizada para a catalogação e feitura de sumários das entrevistas de nosso Programa de História Oral e os verbetes do Dicionário Histórico-Biográfico Brasileiro.

O desenvolvimento do projeto, que conta com a participação de uma ampla equipe de professores, pesquisadores e técnicos do CPDOC, com diferentes formações acadêmicas, tornou evidentes algumas linhas-mestras do tesouro que construíamos ao longo do processo.

Uma das primeiras definições que tivemos que tomar foi em relação ao período cronológico abrangido pelo tesouro. Decidimos incluir o período que vai de 31 de março de 1964, início do golpe militar que depôs João Goulart, até 15 de março de 1985, data de posse do primeiro presidente civil após quase 21 anos de governos militares. Ao fazer isso, não desconhecemos de forma alguma que há na historiografia diferentes periodizações históricas. Também reconhecemos que são usadas várias denominações utilizadas para se referir ao período, a partir de diferentes pressupostos teóricos ou ideológicos. Pretendemos, apenas, delimitar de forma clara um período que servisse de marco cronológico para o tesouro, e que ao mesmo tempo evidenciasse a centralidade política que a instituição militar teve ao longo desse período, no exercício direto do poder: um regime militar que, embora apoiado decisivamente por instituições e personagens civis, teve na instituição militar (representada por sua alta hierarquia) um elemento decisivo na dinâmica política do período.

Essa opção explica a ênfase na dimensão política dos personagens, eventos, instituições e organizações que o tesouro inclui e que evidencia, ressaltamos, características do próprio acervo documental e informacional constituído pelo CPDOC. Neste sentido, alguns elementos foram privilegiados. Associações como a OAB e a ABI, por exemplo, entraram no tesouro, devido à importância de sua atuação política durante o regime militar; não entraram inúmeras outras associações profissionais que existiram durante o regime militar. Alguns intelectuais, artistas e religiosos foram incluídos, ou não, pelos mesmos motivos. Apesar disso, em muitos casos, procuramos incluir alguns elementos que não apareciam nas classificações adotadas em nosso acervo, quer a partir de opções formais – como, p. ex., incluir todos os ministros, governadores, senadores e deputados federais, ou todos os partidos políticos e organizações da esquerda revolucionária que atuaram durante o

período, mesmo que não aparecessem em nossos descritores –, quer a partir do conhecimento histórico acionado pelo grupo que trabalhou na preparação do tesouro.

Não pretendemos, de modo algum, ter feito sempre as melhores opções, nem todas as inclusões pertinentes, mesmo dados os limites dos pressupostos que adotamos. Trata-se, não é demais afirmar, de uma obra sempre em construção, e que será continuamente modificada e adaptada para o uso que faremos desse tesouro no CPDOC. Vale aqui lembrar a tarefa impossível de cobrir todas as possibilidades, tão bem expressa num pequeno texto de Jorge Luís Borges, "Del rigor en la ciência" (publicado em seu livro *El hacedor*, de 1960). O texto fala de um antigo império onde a arte da cartografia era tão perfeita que o mapa do império ocupava toda uma província. Não satisfeitos, os cartógrafos resolveram fazer um mapa ainda mais perfeito, que tivesse o tamanho do império e coincidisse pontualmente com ele. Tal mapa tornou-se, evidentemente, inútil, assim como a própria cartografia. Muitas vezes, ao longo da preparação desse tesouro, tivemos que nos lembrar desse risco.

Feitas essas considerações, resta-nos desejar que o tesouro seja adaptado, aperfeiçoado e mesmo criticado por quem dele fizer uso. Que cada instituição ou pesquisador possa adaptá-lo ou ao menos inspirar-se nele para produzir instrumentos mais adequados aos acervos que forem catalogar – esse é nosso maior desejo, ao torná-lo público. O melhor uso que se pode fazer deste tesouro é justamente estimular a contínua reflexão sobre a prática classificatória, aplicada a contextos históricos específicos. De nossa parte, encaramos institucionalmente a experiência do tesouro como a de uma obra em construção e um aprendizado coletivo. Que ele estimule outras experiências igualmente estimulantes.

REFERÊNCIAS

ANSI/NISO Z.39.19. Guidelines for the construction, format, and management of monolingual controlled vocabularies. Bethesda: NISO Press, 2005. 176 p. Disponível em: <<http://www.niso.org/standards/resources/Z39-19-005.pdf?CFID=2436046&CFTOKEN=81298864>>

BIROU, Alain Dicionário de ciências sociais / Alain Birou ; trad. Alexandre Gaspar e outros.- 5. ed.- Lisboa : Publicações Dom Quixote, 1982.- 454 p.

CAMPOS, Maria Luiza de A.; BRASIL, Maria Irene; COELHO, Beatriz A. S.; BASTOS, Dilza R. *Vocabulário Sistematizado: a experiência da Fundação Casa de Rui Barbosa*. In: INTEGRAR, 2002, São Paulo. INTEGRAR. São Paulo, 2002.

GOMES, Hagar Espanha (Org.). Manual de elaboração de tesouros monolíngues. Brasília: Programa Nacional de Bibliotecas de Instituições de Ensino Superior, 1990.

RANGANATHAN, S. R. Prolegomena to library classification. Bombay, Asia: Publishing House, 1967. 640 p.

O TRATAMENTO DESCRITIVO E TEMÁTICO DE ACERVOS FOTOGRÁFICOS NO PARANÁ

DESCRIPTIVE AND THEMATIC TREATMENT OF PHOTOGRAPHIC COLLECTIONS IN PARANÁ

Ana Cristina de Albuquerque
Ana Carolina Simionato

Resumo: O documento fotográfico presente em Arquivos, Bibliotecas, Museus e demais modalidades de Centros de Documentação é parte de uma tipologia de recursos informacionais com características que devem ser incorporadas, tanto em seu tratamento descritivo, quanto temático. Nesse sentido, o tema dessa investigação parte do seguinte questionamento: como os fundamentos teóricos da Representação e Organização da Informação são utilizados nas instituições que armazenam documentos fotográficos no Estado do Paraná? O objetivo é analisar os fundamentos da Representação e Organização da Informação que atuam em diferentes ambiências informacionais e entender como tais fundamentos são aplicados a documentos fotográficos. Trata-se uma pesquisa ainda em desenvolvimento e sua metodologia é de base exploratória e descritiva, de cunho teórico e prático. Espera-se que os resultados apontem a existência de princípios das Instituições - que sejam específicos a cada campo -, sua relação com a sociedade e com o tipo de documento, além de esclarecer que há analogamente em cada uma das áreas de estudo a necessidade de produzir e reunir conhecimentos sobre o documento fotográfico.

Palavras-chave: Tratamento temático. Tratamento descritivo. Documento fotográfico. Acervo fotográfico – Paraná.

Abstract: The photographic document present in Archives, Libraries, Museums and other Documentation Centres, is part of a typology of information resources with features that should be incorporated both in its descriptive as thematic treatment. Accordingly, the theme of this research is conceptualize from the question: as the theoretical foundations of representation and organization of information are use in the institutions that store photographic documents in Paraná State? The objective is to analyze the fundamentals of Representation and Information Organization that work in different ambiences informational and how these fundamentals are in respect of photographic documents. It search that is under development and its methodology is exploratory and descriptive base of theoretical and practical nature. It is expect that the results point to the existence of the institutions and principles that are specific to each field, and its relationship with society and with the type of document, but also show that there are similarly in each of the study areas the need knowledge about the photographic document.

Keywords: Thematic treatment. Descriptive treatment. Photographic document. Photographic collection - Paraná.

1 INTRODUÇÃO

A necessidade de atenção aos diferentes tipos de recursos informacionais para a preservação e disseminação de suas informações e, conseqüentemente, sua relação com a memória de um lugar ou acontecimento tem sido objeto de inúmeros trabalhos entre os profissionais da informação. Os documentos fotográficos são figuras recorrentes nestas discussões. Eles estão inseridos nas narrativas de investigações, na pesquisa histórica sobre

lugares e fatos, na possibilidade de conhecer, através deles, vestígios do que foi e de como se formou uma cidade, uma família, a organização social de uma nação ou instituição.

Por se referirem a micro aspectos do mundo e a determinados momentos da realidade, o potencial informativo dos documentos fotográficos pode ser alcançado na medida em que sejam contextualizados em sua trama histórica, em seus múltiplos desdobramentos sociais, políticos, culturais, etc. (KOSSOY, 2001). Destarte, Tagg (2005, p. 12, tradução nossa) afirma, a respeito do aparecimento e reconhecimento da fotografia, que este “[...] coincidiu com a variedade de instrumentos e transformações gerados pela natureza da sociedade e, claro, em modos de pensamento, representação e tentativas de agir sobre ele”.

Neste sentido, a obrigação de preservá-los e conhecer sua real situação nas instituições tem caráter urgente, visto que o tratamento de documentos fotográficos, tanto em seus aspectos de conservação, preservação, acondicionamento adequado, quanto em aspectos que se referem ao tratamento temático e descritivo é de extrema importância para a recuperação e acesso a esse tipo de documentação. Portanto, para os documentos fotográficos, as formas de representação temática e descritiva perpassam conceitos que resultam em um aprofundamento e conhecimento dos seus principais fundamentos.

O tratamento de documentos fotográficos insere-se em uma importante atividade que evidencia o papel que da Ciência da Informação na atualidade: o trabalho direto com a informação e o estudo de todos os processos a que essa informação está submetida até que chegue a seus receptores. Desta forma, a presente pesquisa busca entender como os fundamentos teóricos da Representação e Organização da Informação são utilizados nas instituições que armazenam documentos fotográficos no Estado do Paraná.

O objetivo geral é estudar os fundamentos das representações descritiva e temática de documentos fotográficos. Como objetivos específicos, pretende-se levantar as bases teóricas das representações descritiva e temática a fim de estruturar suas definições; verificar como se dá a representação de documentos fotográficos nos acervos selecionados nesta pesquisa; diagnosticar os acervos no âmbito de sua representação descritiva e temática; comparar e confrontar o tratamento temático e descritivo com os manuais clássicos brasileiros que contemplam o tratamento de documentos fotográficos.

Assim, essa investigação se caracteriza por uma metodologia de base exploratória e descritiva, com uma abordagem qualitativa de cunho teórico e prático, tendo como alicerce a literatura disponível que subsidie as novas contribuições acerca do tratamento descritivo e temático dos documentos fotográficos. A forma de coleta dos dados está sendo realizada por

questionários e entrevistas que possibilitarão maior contato com os pesquisadores e profissionais que realizam trabalhos relacionados aos documentos fotográficos.

A contextualização e entendimento do documento fotográfico no âmbito de sua representação temática e descritiva se consolidam por meio de um diagnóstico que está sendo realizado em instituições, no Estado do Paraná, que congreguem em seus acervos estes tipos de documentos. Para tanto, as instituições escolhidas são: Arquivo Público de Londrina, Arquivo Público do Paraná, Museu Histórico de Londrina, Museu Paranaense, Biblioteca Pública do Paraná, Museu da Fotografia da Cidade de Curitiba, Museu da Gravura de Curitiba, Centro de Arte Digital de Curitiba e Centro de Estudos do Movimento de Curitiba.

A pesquisa encontra-se ainda em desenvolvimento e será dividida em duas etapas: o estudo conceitual a respeito da representação descritiva e temática destes registros; e a discussão, que se baseará na compreensão de que todas as normas vigentes em diferentes ambiências informacionais podem englobar objetos singulares como os documentos fotográficos. Esse estudo integra-se ao grupo de pesquisa: ‘Organização e Representação da Informação e do Conhecimento de Recursos Imagéticos (GP-ORICRI)’ da Universidade Estadual de Londrina (UEL).

Continuamente, os resultados obtidos pela coleta serão analisados, a fim de contribuir com a literatura levantada que possibilitarão o confronto com manuais brasileiros que auxiliam no tratamento de recursos imagéticos.

2 CONSIDERAÇÕES SOBRE O DOCUMENTO FOTOGRÁFICO NA ÓPTICA DA REPRESENTAÇÃO TEMÁTICA E DESCRITIVA

Os termos “representação descritiva” e “representação temática” são utilizados com o intuito de evidenciar as diferentes fases de tratamento a que são submetidos os diversos tipos de recursos informacionais, incluindo os documentos fotográficos. Essa diferenciação se refere às etapas que, ao mesmo tempo, são distintas em sua forma de elaboração, mas complementares no sentido da construção de registros para catálogos e bases de dados.

A representação descritiva busca a individualização do documento a fim de que sua representação seja específica. Ou seja, reúne itens por sua semelhança e estabelece relações entre eles (MEY, 1995). A catalogação deve “identificar as possíveis necessidades de informação dos usuários, seguida da criação de uma etiqueta (*tag*), de um texto, ou de um resumo, que vai intermediar o acesso, a identificação e a avaliação do usuário em relação ao recurso original”. (SANTOS, 2013, não paginado). E, portanto, a catalogação facilita que, em

sua busca, o usuário possa identificar, selecionar e recuperar o documento ou o recurso informacional que foi descrito, bem como seu conteúdo, sem dificuldades.

A representação temática, atividade complementar à descritiva, auxilia na recuperação de documentos pois tem a função de expor seus conteúdos. É caracterizada pela atribuição de assuntos aos documentos a partir da classificação, indexação e da elaboração de resumos. De acordo com Dias e Naves (2013, p. 07) “o tratamento temático [...], tem uma forte carga subjetiva, pois, como o nome indica, visa caracterizar o documento do ponto de vista do seu conteúdo. É o que muitos chamarão de assunto do documento, só que esse assunto dependerá muito de quem faz a leitura”.

Além disso, na representação temática há uma lacuna em relação aos documentos fotográficos, pois “No que tange à representação temática, pode-se afirmar que há uma extensa bibliografia que consolida essa etapa para os materiais gráficos, porém no que se refere aos materiais não-gráficos percebe-se lacunas teóricas e metodológicas que necessitam de atenção.” (MAIMONE; GRACIOSO, 2007, p. 01).

Assim, tanto a representação descritiva, quanto a temática se integram caracterizando a socialização de todo o acervo com o usuário, e na medida em que é elaborada uma organização por meio de título, autor, ano e assunto, todos os documentos são disponibilizados com clareza e objetividade, direcionando as representações com a maior especificidade possível e, para que as representações sejam realizadas, é necessário, primeiramente, definir a tipologia do recurso trabalhado de forma a ressaltar as especificidades dos materiais que serão tratados.

Em relação aos instrumentos das representações temática e descritiva, há uma enorme carência de manuais, códigos e padrões que fundamentem as características intrínsecas e extrínsecas do recurso imagético, especialmente do documento fotográfico. Por isso, aponta-se a necessidade de estudar a essência representacional de que este tipo de documento necessita, para que de acordo com o usuário especializado seja possível a recuperação e acesso ao recurso.

Neste sentido, Simionato e Santos (2013, p. 3) explicam que o registro do recurso imagético em geral, independentemente de seu ambiente ser analógico ou digital, demonstra que as características “[...] não são abordadas e muitas vezes os atributos são valorados por características textuais apenas contextualizadas”. Dessa forma, os manuais existentes representam o documento fotográfico à mesma maneira do livro.

Neste contexto, entende-se que os manuais não oferecem o suporte necessário para que estes documentos específicos sejam devidamente representados. Como também, os

manuais não oferecem a representação e organização necessária de acordo com os acervos de documentos fotográficos que vão se formando, sejam estes de natureza institucional ou pessoal, que refletem o pensamento de uma época.

Há de se considerar que as formas de organizar e representar documentos devem ser mantidas e respeitadas pelas instituições que os recolhem, a fim de “[...] não apagar os traços da sua organicidade, traduzida no modo como ela foi acumulada, reunida e, ainda, naquilo que foi intercambiável nessa reunião [...]” (GONÇALVES; MARCONDES, 2005, p. 263).

3 O DOCUMENTO FOTOGRÁFICO EM BIBLIOTECAS, ARQUIVOS, MUSEUS E CENTROS DE DOCUMENTAÇÃO

Investigar os parâmetros teóricos que cercam a área da Ciência da Informação, assim como definir conceitos, exige examinar as fontes onde possam ser apreendidas informações que subsidiem discussões acerca do que foi sugerido. Esse parâmetro torna-se um desafio especialmente pela dificuldade em reunir bibliografias, por vezes dispersas, e compor discussões que apenas recentemente vêm sendo pensadas com maior frequência no âmbito da Ciência da Informação.

Os Arquivos, permeados pelos princípios de proveniência e de integridade, são elucidados ao contato com o documento fotográfico no momento de aquisição de um fundo, e em muitas vezes seu tratamento descritivo e temático é designado em conjunto a outras tipologias documentais. Heredia Herrera (1993) aponta que a principal observação é a vinculação institucional sendo o diferencial para os diversos suportes, e por isso há diversas questões sobre o documento fotográfico que necessitam aprofundamento nos Arquivos, entre as quais: terminologia, conservação, formas de agrupamento, o tratamento que transita entre a dupla vertente de classificação e de análise.

Quando em Bibliotecas, o documento fotográfico nem sempre é incorporado em catálogos ou nas listas de registros e, por isso, poucos profissionais conhecem a capacidade informacional desse tipo de recurso. O que gera pouca literatura sobre documentos fotográficos em Bibliotecas. Mas deve-se destacar que em algumas grandes bibliotecas esse tipo de recurso recebe o tratamento adequado, como na *Library of Congress* (LC).

Por fim, nos Museus e nos Centros de Documentação, o documento fotográfico é encarado como coleção e, por isso, “Os princípios de organização dos quais depende a coleção sistemática pressupõe uma realidade observável e operações racionais que permitam a seleção deste ou daquele objeto do mundo material para ser considerado, estudado, preservado.” (LIMA; CARVALHO, 2000, p. 19).

Portanto, ao investigar as diferentes ambiências informacionais entre Arquivos, Bibliotecas, Museus e Centros de Documentação, fica claro que há pontos que coincidem por algumas perspectivas científicas comuns. De forma que, independentemente da Unidade Informacional, nem sempre se perceberá os mesmos traços constitutivos, sendo diferentes as funções no que dizem respeito ao documento fotográfico (ALBUQUERQUE, 2006).

4 CONSIDERAÇÕES PARCIAIS

O tratamento de documentos fotográficos pelos diferentes campos é sistematizado a partir de elementos que são ordenados a fim de representar o seu contexto, cujo principal objetivo é o acesso à informação.

Em quaisquer aspectos que se encontrem, a partir do momento em que integram uma instituição, os documentos fotográficos necessitam de métodos e tratamentos que façam refletir, da forma mais objetiva possível, suas informações, por vezes claras e facilmente perceptíveis, por vezes, devido a sua contextualização e produção, de difícil acesso aos profissionais envolvidos em seu tratamento.

Neste sentido, os tratamentos temático e descritivo são atividades complexas que congregam os princípios da representação e da prática e, por esse motivo, requerem aprofundamento e políticas que especifiquem as características relevantes para o seu tratamento sem ignorar que este tipo de documento não pode ser tratado conjuntamente a outro tipo de acervo.

REFERÊNCIAS

ALBUQUERQUE, A. C. **Catálogo e descrição de documentos fotográficos em bibliotecas e arquivos: uma aproximação comparativa dos códigos AACR2 e ISAD (G)**. 2006. 188f. Dissertação (Mestrado em Ciência da Informação)–Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2006.

DIAS, E. W.; NAVES, M. M. L. **Análise de assunto: teoria e prática**. 2. ed. Brasília, DF: Briquet Lemos, 2013.

GONÇALVES, C. D.; MARCONDES, M. A coleção fotográfica V-8. In: **Anais do Museu Paulista: história e cultura material**, São Paulo, v. 13, n. 1, p. 253-269. 2005. Disponível em: <<http://www.redalyc.org/articulo.oa?id=27313109>>. Acesso em: 22 jul. 2014.

HEREDIA HERRERA, A. La fotografía e los archivos. In: FORO IBEROAMERICANO DE LA RÁBIDA. Jornadas Archivísticas, 2., 1993. **Anais...** Huelva: Diputación Provincial, 1993.

KOSSOY, B. **Fotografia e história**. São Paulo: Ateliê Editorial, 2001.

LIMA, S. F. de; CARVALHO, V. C. Cultura material e coleção em um museu de história: as formas espontâneas de transcendência do privado. In: FIGUEIREDO, B. G.; VIDAL, D. G.

(Org.). **Museus**: dos gabinetes de curiosidades ao museu moderno. Belo Horizonte, MG: Argvmentvm; Brasília, DF: CNPq, 2005.

MAIMONE, G. D.; GRACIOSO, L. S. Representação temática de imagens: perspectivas metodológicas. **Informação & Informação**, v. 12, n. 1, jan./jun. 2007. Disponível em: <<http://www.uel.br/revistas/uel/index.php/informacao/article/view/1760>>. Acesso em: 20 jul. 2014.

MEY, E. S. A. **Introdução à catalogação**. Brasília: Briquet de Lemos, 1995.

SANTOS, P. L. V. A. C. Catalogação, formas de representação e construções mentais. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 6, n. 1, 2013. Disponível em: <<http://inseer.ibict.br/ancib/index.php/tpbci/article/view/100>>. Acesso em: 29 jul. 2014.

SIMIONATO, A. C.; SANTOS, P. L. V. A. C. Descrição de recursos imagéticos digitais: apresentação de um modelo conceitual. **Anales de Documentación** (Internet), v. 16, p. 1-17, 2013. Disponível em: <<http://revistas.um.es/analesdoc/article/view/179261>>. Acesso em: 20 jul. 2014.

TAGG, J. **El peso de la representación**: ensayos sobre fotografías e historias. Barcelona: Gustavo Gilli, 2005.

EXAUSTIVIDADE, PRECISÃO E CONSISTÊNCIA EM INDEXAÇÃO: SISTEMATIZAÇÃO CONCEITUAL

EXHAUSTIVITY, ACCURACY AND CONSISTENCY IN SUBJECT INDEXING: CONCEPTUAL SYSTEMATIZATION

Luciana Beatriz Piovezan
Mariângela Spotti Lopes Fujita

Resumo: Considerando-se a indexação enquanto procedimento de organização e representação do conhecimento, teve como propósito apresentar uma sistematização conceitual dos elementos exaustividade, precisão e consistência, enquanto elementos constituintes da política de indexação, após a constatação da divergência na forma como são entendidos por diferentes autores. Foi realizada uma investigação conceitual destes elementos na literatura em que se aborda a indexação como processo de organização e representação do conhecimento em Ciência da Informação. Foram discutidos os conceitos propostos pela literatura para os elementos exaustividade, precisão e consistência da indexação, de modo a perceber os diversos aspectos destes elementos tratados pela literatura. A análise evidenciou a relação entre a indexação e a recuperação da informação como processos que se complementam. Apresenta nas considerações finais uma definição que procura ressaltar os diversos aspectos identificados na literatura para os elementos analisados.

Palavras-chave: Indexação de assuntos. Política de indexação. Consistência. Exaustividade. Precisão.

Abstract: Considering subject indexing as a knowledge organization and representation procedure, we aimed to present a conceptual systematization of the elements exhaustivity, accuracy and consistency, as components of subject indexing policies, after the recognition of disagreements in how they are understood by different authors. In order to achieve that, a conceptual investigation of these fundamental elements was conducted in the literature that addresses the subject indexing as a process of knowledge organization and representation in Information Science. The concepts were discussed in order to understand the different aspects of these elements covered in the literature. The analysis demonstrated the relationship between subject indexing and information retrieval as processes that complement each other. Thus, presents in the final considerations a definition that seeks to highlight the different aspects identified in the literature for the elements analyzed.

Keywords: Subject indexing. Subject indexing policy. Consistency. Exhaustivity. Accuracy.

1 INTRODUÇÃO

A indexação como processo de organização e representação do conhecimento integra-se aos estudos da informação em razão da relação de causa e efeito entre representação e recuperação da informação. Os efeitos do processo de indexação são continuamente investigados mediante os elementos presentes tanto no processo de indexação quanto na recuperação da informação. Os elementos do processo de indexação são indicados na literatura de forma divergente entre os autores e os que se destacam dentre vários são exaustividade, especificidade, revocação e precisão porque são indicadores de avaliação para os quais existem formulas matemáticas de aferição. Este trabalho dedicou-se à investigação

conceitual dos elementos de política de indexação exaustividade, precisão e consistência no contexto da política de indexação mediante estudo na literatura fundamental em que se aborda a indexação como processo de organização e representação do conhecimento em Ciência da Informação com o objetivo de apresentar sistematização conceitual dos elementos e distinguir suas finalidades e funções.

2 POLÍTICA DE INDEXAÇÃO

Segundo Carneiro (1985) as políticas de indexação são orientações para a tomada de decisões no planejamento de sistemas de recuperação da informação, de forma a garantir que qualquer informação seja provida ao usuário no momento preciso, da maneira mais eficiente e econômica possível. Rubi afirma ainda que a política de indexação tem por metas “harmonizar o acesso por assunto em seu catálogo e explicitar as decisões da indexação para os catalogadores, para os profissionais que importam seus registros e para os usuários que fazem buscas por assunto” (2008, p. 51).

Os objetivos do desenvolvimento de uma política de indexação são definir as variáveis que afetam o desempenho do sistema de informação, estabelecer o critério e princípios que guiarão a tomada de decisões para tornar o sistema mais eficiente, a racionalização dos processos e a consistência das operações neles envolvidas (CARNEIRO, 1985).

Para Gil Leiva a política de indexação congrega não apenas o modo com a indexação será realizada, mas também representa o esforço de consolidar, sistematizar e esclarecer por meio de manuais os processos realizados na indexação (GIL LEIVA, 2008). Dessa forma, compreendemos que a política de indexação deve fazer parte do planejamento do sistema de informação “[...] de modo a compor-se como um conjunto de procedimentos, materiais, normas e técnicas orientadas por decisões que refletem a prática e princípios da cultura organizacional” (FUJITA; GIL LEIVA, 2009, p. 156).

Carneiro (1985) estabelece que para a construção de uma política de indexação é necessário identificar as características do usuário, suas áreas de interesse, nível educacional, experiência e atividades que exercem; identificar as características da literatura que será incorporada ao sistema; as características das solicitações dos usuários; os recursos humanos e financeiros envolvidos para criação e manutenção do sistema; e os equipamentos disponíveis (CARNEIRO, 1985, p. 165).

Os elementos constituintes de uma política de indexação para a autora são a cobertura de assuntos, seleção e aquisição de documentos-fonte, o processo de indexação (composto por nível de exaustividade, nível de especificidade, escolha da linguagem e capacidade de

revocação e precisão do sistema), estratégia de busca, tempo de resposta do sistema, formato de saída dos dados e avaliação do sistema (CARNEIRO, 1985, p. 231).

Em estudos mais recentes estes elementos receberam atualizações, dentre as quais destacamos a contribuição de Guimarães (2004) que apontou, tendo por base o trabalho de Foskett (1973), a necessidade de incluir os elementos capacidade de consulta a esmo (*browsing*) e garantia literária, o primeiro relacionando-se a usabilidade do sistema de busca e o segundo relacionado à utilização da linguagem.

Também destacamos a contribuição realizada por Rubi (2008) que incluiu como elementos de política de indexação a formação do indexador e manual de indexação e, apoiada em Olson e Boll (2001), incluiu os elementos adequação, exaustividade, especificidade e consistência. Rubi destaca que embora Olson e Boll (2001) assim como outros, não considerem tais características como elementos de política de indexação, o fato de serem variáveis que afetam a recuperação da informação os coloca como pertinentes à política de indexação.

Para este estudo foram selecionados os elementos exaustividade, precisão e consistência para um maior detalhamento, dadas às suas características particulares. Em primeiro lugar, tais elementos influenciam diretamente na recuperação da informação garantindo ou fazendo com que seja perdida a confiabilidade do sistema, e segundo, são variáveis que podem ser medidas por índices.

3 EXAUSTIVIDADE, PRECISÃO E CONSISTÊNCIA NA INDEXAÇÃO

Em estudos de política de indexação se sobressai a preocupação com os elementos exaustividade, especificidade e consistência da indexação, tais elementos conferem ao sistema de indexação um equilíbrio entre revocação e precisão na recuperação da informação e por sua observação é possível avaliar a flexibilidade deste sistema frente às necessidade de exaustividade e precisão demandada pela comunidade usuária (GIL URDICIAIN, 1997). No intuito de apresentar uma sistematização destes elementos, realizamos uma análise comparativa das definições encontradas na literatura de Ciência da Informação para os termos em questão.

3.1 Exaustividade na indexação

Para o elemento exaustividade foram localizadas definições propostas por Lancaster (2004), Foskett (1973), Olson e Boll (2001), Gil Leiva (1999), Gil Urdiciain (1997) e Rubi (2008), assim como Carneiro (1985) que se apoiou em Lancaster (1968). Para melhor visualização das definições obtidas apresenta-se o QUADRO 1.

Podemos observar que as definições propostas por Lancaster (2004) e Rubi (2008) tratam a exaustividade como uma medida que quantifica os termos atribuídos aos documentos para a representação de seu conteúdo.

QUADRO 1 – Definições do elemento exaustividade

LANCASTER	“emprego de termos em número suficiente para abranger o conteúdo temático principal do documento. Quanto mais termos forem utilizados para indexar um documento, mais acessível ele se tornará e, provavelmente, mais vezes será recuperado” (LANCASTER, 2004, p. 27).
FOSKETT	“a extensão com que analisamos um dado documento, a fim de estabelecer exatamente qual o conteúdo temático que temos de especificar” (FOSKETT, 1973, p. 13).
CARNEIRO	“medida de extensão em que todos os assuntos discutidos em um certo documento são reconhecidos na operação de indexação e traduzidos na linguagem do sistema” Lancaster (1968).
OLSON; BOLL	“é o número de conceitos representados no registro bibliográfico ou a abrangência do assunto coberto” (2001, p. 91).
GIL LEIVA	“A seleção ou assinalação de todos os conceitos essenciais, explícitos ou implícitos no documento ou na pergunta documental” (1999, p. 26).
GIL URDICIAIN	“a exaustividade mede a capacidade do sistema para recuperar documentos úteis” (1997, p. 277).
RUBI	“O número de termos utilizados para descrever o documento diz respeito à exaustividade que por sua vez está relacionada à revocação e à precisão do sistema de recuperação” (2008, p. 35).

Fonte: Elaboração própria.

As definições propostas por Olson e Boll (2001), Gil Leiva (1999) e Carneiro (1985) apresentam a exaustividade como relacionada à quantidade de conceitos que representam o conteúdo do documento, ou seja, com a seleção de termos que compreendam todos os conceitos presentes no documento e não somente o número absoluto de termos atribuídos.

A definição proposta por Foskett (1973) expõe o aspecto da análise exaustiva do documento a fim de conhecer a completude de seu conteúdo temático. Já a autora Gil Urdiciain (1997) apresenta a exaustividade observada pela perspectiva da recuperação da informação em sistemas construídos para esta finalidade, como a qualidade do sistema para recuperar todos os documentos presentes que respondam à questão de busca.

3.2 Precisão no processo de indexação

No QUADRO 2 apresentamos as definições localizadas para o elemento precisão, as quais foram propostas por Foskett (1973), Carneiro (1985), Olson e Boll (2001), Svenonious (2000), Hudon (2009) e Gil Urdiciain (1997).

Observando as definições identificadas na literatura para o elemento precisão é possível identificar que se situam entre dois aspectos: o da exatidão na determinação do assunto de um documento (FOSKETT, 1973; OLSON; BOLL, 2001), e o da qualidade dos sistemas de recuperação da informação na identificação de documentos relevantes (CARNEIRO, 1985; SVENONIOUS, 2000; HUDON, 2009; GIL URDICIAIN, 1997).

QUADRO 2 – Definições do elemento precisão

FOSKETT	“a extensão em que o sistema nos permite ser precisos ao especificarmos um documento que estejamos processando” (1973, p. 12)
CARNEIRO	“capacidade do sistema em impedir a recuperação dos documentos não-relevantes” (1985, p. 234)
OLSON; BOLL	“a habilidade do catalogador de identificar quais conceitos representar e a habilidade do catalogador para traduzir estes conceitos em termos de um vocabulário controlado” (2001, p. 88)
SVENONIOUS	“o grau com o qual uma linguagem de assunto é capaz de anular a seleção de documentos irrelevantes na recuperação” (2000, p. 189)
HUDON	“medida de desempenho de um sistema de informação que define a proporção de documentos relevantes recuperados em comparação com o conjunto que consiste de todos os documentos fornecidos em resposta a um pedido” (2009, p. 259)
GIL URDICIAIN	“a precisão mede a habilidade [do sistema] de rejeitar material não relevante” (1997, p. 277)

Fonte: Elaboração própria.

Estes aspectos refletem a relação de causa e efeito em indexação e busca de informação. O efeito na recuperação da informação de uma indexação que atende a precisão é a capacidade do sistema de identificar documentos que atendem à questão de busca e evitar documentos que não tem pertinência com a questão, rejeitando documentos irrelevantes para a busca.

3.3 Consistência da indexação

Para compor a análise do elemento consistência foram localizadas definições propostas por Lancaster (2004), Olson e Boll (2001), Zunde e Dexter (1969), Rolling (1981), Moreiro González (2004), Hudon (2009) e Gil Urdiciain (1997). Uma visualização das definições obtidas apresenta-se no Quadro 3.

Por outro lado, Rolling (1981) e Hudon (2009) fazem menção à consistência como a utilização de termos similares na representação documental, podendo inferir-se daí que tais termos possam não ser necessariamente os mesmos, mas possuírem equivalência. As definições propostas por Olson e Boll (2001) e Moreiro González (2004) apresentam a

consistência relativa à igualdade na realização da análise conceitual do documento e também da tradução em termos.

QUADRO 3 – Definições do elemento consistência

LANCASTER	“refere-se à extensão com que há concordância quanto aos termos a serem usados para indexar um documento” (2004, p. 68)
OLSON; BOLL	“a consistência requer que itens sobre o mesmo assunto sejam conceitualmente analisados e traduzidos da mesma forma” (2001, p. 99)
ZUNDE; DEXTER	“o grau de concordância na representação da informação essencial de um documento por meio de um conjunto de termos de indexação selecionados por cada um dos indexadores de um grupo” (1969, p. 259)
ROLLING	“[...] manifesta-se na similaridade dos termos de indexação atribuídos a um dado documento por diferentes indexadores” (1981, p. 69)
MOREIRO GONZÁLEZ	“busca que um conceito ou tema apareça expresso sempre da mesma forma” (2004, p. 51)
HUDON	“medida de similaridade dos resultados obtidos quando vários indexadores são chamados para representar o conteúdo de um documento ou vários documentos sobre o mesmo assunto” (2009, p. 259)
GIL URDICIAIN	“a utilização sempre dos mesmos termos no processo [de indexação]” (1997, p. 390)

Fonte: Elaboração própria.

4 CONSIDERAÇÕES FINAIS

Neste trabalho pretendemos apresentar sistematização conceitual dos elementos exaustividade, precisão e consistência e distinguir suas finalidades e funções. Tais elementos têm como objetivo equilibrar as respostas do sistema, colaborando para uma melhor recuperação da informação. A literatura evidencia a relação destes elementos entre si e com o elemento especificidade. A determinação de critérios de exaustividade, precisão, assim como de especificidade auxiliam a prática da indexação de modo a aumentar o seu grau de consistência.

Dada a análise realizada, podemos inferir que a literatura de Ciência da Informação entende a exaustividade como a extensão da análise documental para identificar com completude os conceitos representativos de um documento e a seleção de termos que abrangem todos os conceitos importantes identificados durante a análise, de modo que o documento seja recuperado em questões de busca para as quais seja útil.

Do mesmo modo, a precisão refere-se à profundidade da análise documental para identificar e traduzir com exatidão o potencial informativo de um documento de modo que o sistema de recuperação da informação seja capaz de identificar os documentos que atendem à questão de busca e evitar documentos que não tenham pertinência com a questão.

Podemos entender que a consistência é percebida como o grau de concordância com o qual documentos são analisados conceitualmente e traduzidos em termos de indexação idênticos ou terminologicamente equivalentes.

REFERÊNCIAS

CARNEIRO, M. V. Diretrizes para uma política de indexação. **Revista da escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 14, n. 2, p. 221-241, set. 1985.

FOSKETT, A. C. **A abordagem temática da informação**. São Paulo: Polígono; Brasília: UnB, 1973.

FUJITA, M. S. L; GIL LEIVA, I. Políticas de indexação na América Latina. In: GARCÍA MARCO, F. J. **IBERSID 2009**. Avances y perspectivas en sistemas de información y documentación, Espanha, 2009.

GIL LEIVA, I. La automatización de La indización de documentos. Gijón: Trea, 1999.

GIL LEIVA, I. **Manual de indización**. Gijón: Trea, 2008.

GIL URDICIAIN, B. **Evolución histórica de los tesauros españoles y análisis de su rendimiento en el proceso de recuperación de información**. 1997. 429f. Tese (Doctorado en Ciencias de la Información) – Facultad de Ciencias de la Información, Universidad Complutense de Madrid, Madrid, 1997.

GUIMARÃES, J. A. C. As políticas de indexação como elemento para a gestão do conhecimento nas organizações. In: VIDOTTI, S. A. G. **Tecnologias e conteúdos informacionais**. São Paulo: Polis, 2004.

HUDON, M. Guide pratique pour l'elaboration d'un thesaurus documentaire. Québec: Asted, 2009.

LANCASTER, F. W. **Information retrieval systems: characteristics, testing and evaluation**. New York: Wiley, 1968.

LANCASTER, F. W. **Indexação e resumos: teoria e prática**. 2. ed. Brasília: Briquet de Lemos, 2004.

MOREIRO GONZÁLEZ, J. A. **El contenido de los documentos textuales: su análisis y representación mediante el lenguaje natural**. Gijón: Trea, 2004.

OLSON, H. A.; BOLL, J. J. **Subject analysis in online catalogs**. 2. ed. Englewood, CO: Libraries Unlimited, 2001.

ROLLING, L. Indexing consistency, quality and efficiency. **Information processing & Management**, v. 17, n. 2, p. 69-76, 1981.

RUBI, M. P. **Política de indexação para construção de catálogos coletivos em bibliotecas universitárias**. 2008. 169 f. Tese (Doutorado em Ciência da Informação) Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2008.

RUBI, M. P.; FUJITA, M. S. L. Política de indexação na catalogação de assunto em bibliotecas universitárias: a visão sociocognitiva da atuação profissional com protocolo

verbal. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v.7, n. 2, p. 118-150, jan./jun. 2010.

SVENONIUS, E. The intellectual foundation of information organization. Cambridge: MIT Press, 2000.

ZUNDE, P.; DEXTER, M. E. Indexing consistency and quality, **American Documentation**, v. 20, n. 3, p. 259-267, 1969.

A ARQUITETURA DE METADADOS DE PROCESSOS JUDICIAIS CÍVEIS

THE METADATA ARCHITECTURE OF LAWSUITS

Eduardo Watanabe

Resumo: O presente trabalho busca apresentar a importância da arquitetura de metadados para a adequada organização e recuperação da informação. Primeiro é feito um panorama do volume excessivo de processos judiciais em tramitação no Brasil para depois discorrer sobre as possibilidades na utilização adequada dos metadados em auxílio à tramitação e julgamento mais ágeis com o objetivo de reduzir a taxa de congestionamento existente. Em seguida, são feitas considerações iniciais sobre os metadados e a arquitetura instituída pelo Conselho Nacional de Justiça. São depois abordados os metadados existentes para assunto, classe e tramitações dos processos judiciais, bem como para o resultado de decisões judiciais. Foi observada a necessidade de ajustes nas tabelas para que a informação possa ser representada com maior precisão, em especial no que se refere ao assunto processual. A utilização de metadados para a recuperação de informação de processos judiciais ainda não foi objeto de estudos no Brasil, o que abre muito campo para futuras pesquisas no assunto.

Palavras-chave: Arquitetura de metadados. Organização da informação. Recuperação da informação. Processos judiciais.

Abstract: This study aims to present the importance of metadata architecture for the suitable organization and retrieval of information. First an overview is made of the excessive volume of lawsuits pending in Brazil and then discuss the possibilities in the appropriate use of metadata to aid in processing and more agile trial aiming to reduce the rate of existing congestion. Then, initial considerations on the metadata and the architecture established by the National Council of Justice are made. Are then discussed the existing metadata for subject, class and formalities of the lawsuits as well as for the result of judicial decisions. The need for adjustments were observed in the tables so that the information can be represented with greater accuracy, especially in regard to the procedural issue. The use of metadata for information retrieval lawsuits has not yet been the subject of studies in Brazil, which opens much scope for future research on the subject.

Keywords: Architecture metadata. Organization of information. Information retrieval. Lawsuits.

1 INTRODUÇÃO

A Ciência da Informação no Brasil tem fornecido diversas contribuições para o desenvolvimento do Direito. O presente artigo busca analisar uma área ainda inexplorada, que é a da utilização de metadados na organização e recuperação de informações do assunto, classe e tramitações dos processos judiciais cíveis, além dos resultados de decisões judiciais.

O Poder Judiciário no Brasil tramita cada vez mais processos judiciais, que saltaram de 83,4 milhões em 2009 para 92,2 milhões em 2012, conforme informações consolidadas pelo Conselho Nacional de Justiça (BRASIL, 2013). Não obstante a taxa de congestionamento total do Poder Judiciário estar em queda, ela ainda estava em 70% no ano

de 2012, o que significa que, a cada 100 processos tramitados no ano, apenas 30 foram encerrados (BRASIL, 2013).

O presente trabalho busca responder à seguinte pergunta: os metadados utilizados atualmente pelo Poder Judiciário são adequados para que todo o volume de informação possa ser organizado de forma adequada em proveito de defesas e julgamentos mais ágeis?

2 DOS PROCESSOS JUDICIAIS CÍVEIS

Processo judicial é um conjunto de documentos produzidos pelas partes e órgãos judiciários que tem como objetivo solucionar uma lide (conflito de interesses). No Brasil, cabe ao Poder Judiciário conduzir o trâmite dos processos judiciais, tudo de acordo com as prescrições contidas em leis federais.

O processo judicial cível é iniciado em regra por um documento elaborado pela parte autora chamado petição inicial, que é direcionado a um juiz (art. 295 do Código de Processo Civil). Por regra, o juiz determina a citação da parte ré, para que este venha se defender em juízo. A partir desse momento se aperfeiçoa a relação jurídica processual, que é triangular: parte autora, juiz e parte ré.

Todo juiz é auxiliado por uma secretaria judiciária, cujos membros são responsáveis pela organização do processo, o que inclui a juntada de petições e documentos recebidos das partes e os atos praticados pelo juiz e pelos próprios servidores da secretaria (art. 141 do Código de Processo Civil).

A informatização da justiça pode ser dividida em duas situações: sistemas para tramitação de processos físicos; sistemas para tramitação, acesso e intimações de processos eletrônicos. Todos os tribunais brasileiros utilizam pelo menos o sistema de tramitação de processos, de modo que temos em todos os processos judiciais em andamento o controle dos dados cadastrais, das tramitações e dos resultados de decisões judiciais.

Quanto ao processo judicial eletrônico ele tem avançado, ainda que de forma heterogênea, com participação nos casos novos de 13% na Justiça Estadual, 15,2% nas Varas da Justiça do Trabalho e 65% na Justiça Federal (CNJ, 2013). Contudo, é significativo notar que o atual formato dos processos judiciais eletrônicos não representa uma evolução na organização da informação contida nos documentos que fazem parte do processo, pois o “processo eletrônico” no momento é sinônimo de processo digitalizado ou arquivo em formato .PDF. Ou seja, os usuários precisam ainda fazer a “constatação visual” dos documentos do processo eletrônico, a leitura das imagens digitais, de forma muito parecida com a recuperação de informação feita nos processos físicos (PEREIRA, 2009).

Os processos eletrônicos utilizam os mesmos padrões de metadados utilizados para os processos físicos, por isso podemos afirmar que ele ainda não trouxe para os usuários nenhuma novidade na apresentação das informações.

3 DOS METADADOS DE PROCESSOS JUDICIAIS CÍVEIS

O Conselho Nacional de Justiça (CNJ), órgão responsável pelo controle da atuação administrativa e financeira do Poder Judiciário, firmou com outros órgãos do Judiciário, a Advocacia-Geral da União e a Procuradoria-Geral da República o compromisso de adoção de um modelo nacional de interoperabilidade (MNI), que contém padrões para intercâmbio de informações de processos judiciais por meio de sistemas processuais (CNJ, 2009).

Para fins do presente estudo, iremos analisar o MNI vigente no que se refere ao tipo Processo Judicial (classe), tipo Assunto Processual e tipo Movimento Processual (tramitação e resultado de decisão judicial) (CNJ, 2012). Para cada um dos tipos, o CNJ elaborou e mantém uma tabela respectiva (CNJ, 2014), cujo processo de implantação está em andamento (CNJ, 2010). Como as tabelas disponibilizadas estão por Justiça e Grau, iremos utilizar a tabela referente à Justiça Federal e ao Superior Tribunal de Justiça (STJ) por tratarem de processos judiciais cíveis.

A utilização de metadados em dados digitais tem provocado a utilização do termo “metadados” em situações muito discrepantes (CAMPOS, 2007). No presente artigo utilizamos o conceito de metadados de Gilliland (2008, p. 2), que foca mais na sua finalidade ao conceitua-los como: “a soma total do que pode ser dito sobre algum *objeto informacional* em algum nível de agregação”.

As informações relativas aos processos judiciais podem ser assim divididas: informações sobre o conteúdo do processo (classe e assunto); informações sobre a alteração de responsabilidades e localização do processo (tramitação); informações sobre o teor de documento do processo (resultado de decisão judicial). Na classificação de Gilliland (2008), os metadados analisados são criados manualmente por humanos, são de natureza especializada, de longo prazo, pois utilizados para garantir o seu acesso e uso, com estrutura e semântica controladas.

Em levantamento bibliográfico, Campos (2007) identificou dez categorias funcionais dos metadados: Descritividade, Padronização, Interoperabilidade, Modularidade, Reflexividade, Automatização, Administração e Preservação. Para os metadados analisados, identificamos a ênfase em Descritividade e Padronização com nítido direcionamento futuro para a Interoperabilidade.

3.1 Classe processual

A tabela de classes processuais da Justiça Federal de 1º grau possui 267 códigos, sendo cada um discriminado com as seguintes colunas: nome da classe, código, código pai, dispositivo legal, artigo, sigla, alterações e glossário (CNJ, 2014).

Os códigos são definidos a partir do fundamento em dispositivo legal, o que vincula o metadados ao tipo de procedimento. Em geral a adoção do rito como critério de elaboração da tabela não provoca perda de informação. Podemos citar os casos das classes “Reintegração/Manutenção de posse”, “Usucapião” e “Renovatória de locação”. Nessas situações fica expresso que o pedido da parte autora está evidenciado no próprio nome da classe processual. E além do pedido o próprio assunto do processo fica bem expresso.

Porém, há classes processuais que não coincidem com o pedido da parte autora devido ao seu grau de generalidade. Na classificação quinária de Miranda (2002), o pedido do autor veicula uma tutela declaratória, condenatória, constitutiva, mandamentais ou executivas *lato sensu*. Na classe Procedimento Ordinário, código 7, está prevista nos artigos 272 e 274 do Código de Processo Civil e consta no glossário a orientação para utilizar o código como residual, o que deve concentrar um volume muito grande de informação sem o necessário detalhamento.

O mesmo se aplica a diversas outras classes processuais na tabela como o Procedimento Sumário, a Ação Civil Pública, o Mandado de Segurança dentre outras.

3.2 Assunto processual

Na tabela de assuntos processuais da Justiça Federal de 1º grau está o maior número de códigos com 1.754, sendo cada um discriminado com as seguintes colunas: nome do assunto, código, código pai, dispositivo legal, artigo, sigla, alterações e glossário (CNJ, 2014).

Ao contrário da tabela de classes processuais, a de assuntos contém diversos códigos sem a expressa remissão a dispositivo ou artigo legal. Isso se explica pelo fato de que há uma produção legislativa muito volumosa no Brasil, que abrange os níveis federal, distrital, estadual e municipal, sendo que isso não ocorre na mesma proporção com as leis processuais. Devido a isso, a tabela de assuntos é a que possui mais perspectivas de crescimento ao mesmo tempo em que tem mais lacunas.

Em comparação, a Classificação Decimal de Direito (CDD) dispõe de cerca de 3.000 códigos para classificação, sendo que não se constata harmonia entre as duas formas de organização dos assuntos. Ainda que se tratem de objetos de aplicação diferentes, a CDD

para acervos bibliográficos e a tabela de assuntos para documentos arquivísticos, é natural que se pudesse esperar uma aproximação maior entre ambas.

Andreta (2004) propõe a extensão da CDD para que possa abranger novas informações que surgiram com a evolução dos conhecimentos na área jurídica, o que revela o permanente desafio de aperfeiçoamento dos códigos de classificação por assunto no Direito.

Essas dificuldades de indexar o assunto do documento são expostas por Guedes (2009) em sua dissertação sobre o *aboutness*, também conhecido como tematicidade ou atinência. Ele conclui que os estudos no tema são ainda muito tímidos, não obstante a sua inegável importância para o desenvolvimento da Ciência da Informação. Faz-se necessário uma pesquisa de campo para avaliar se os usuários estão satisfeitos com as opções de classificação de assunto que lhe são fornecidas e a qualidade da própria indexação que é feita.

3.3 Tramitações

As tramitações estão na tabela de movimentações da Justiça Federal de 1º grau somente quanto aos itens de responsabilidade de serventuários, o que totaliza 61 códigos, sendo cada um discriminado com as seguintes colunas: nome do assunto, código, código pai, complemento, movimento, visibilidade externa, dispositivo legal, artigo, sigla, alterações e glossário (CNJ, 2014). Os códigos estão agrupados por tipo de função do servidor: arquivista, contador, distribuidor, escrivão/diretor de secretaria/secretário judiciário e oficial de justiça.

A informação é útil para os usuários operacionais terem a informação sobre a situação ou ocorrências no processo referentes às atividades de apoio (juiz, partes do processo, servidores da secretaria). Considerando que estão cada vez mais presentes estudos e práticas de aperfeiçoar a gestão judiciária (SILVA, 2012; SILVA *et. ali*, 2010; MACIEIRA; MARANHÃO, 2010; NOGUEIRA, 2010; FREIRE, 2004), as informações podem ser agregadas para compor indicadores que poderão auxiliar no auxílio à gestão judiciária, o que tem como principais interessados a Corregedoria dos Tribunais, o juiz e o diretor de secretaria.

3.4 Resultados das decisões judiciais

Os resultados das decisões judiciais estão na tabela de movimentações da Justiça Federal de 1º grau apenas para os itens de responsabilidade de juízes, o que totaliza 229 códigos (CNJ, 2014). Os códigos estão agrupados por tipo ato do juiz, previstos no Código de Processo Civil e que se referem cada uma a tipologia documental própria: decisão interlocutória, despacho e sentença.

No segundo e terceiro níveis de detalhamento estão o tipo de providência ou o resultado, e em um eventual quarto nível fornece uma informação mais detalhada. Os seguintes exemplos podem ilustrar melhor o exposto: julgamento por meio de sentença (1º nível), com resolução do mérito (2º nível), concessão (3º nível) de *habeas corpus* (4º nível); decisão (1º nível), concessão (2º nível) e liminar (3º nível).

Os metadados de resultados de decisões judiciais são de muita utilidade para a adequada gestão na advocacia pública e em escritórios de advocacia (WATANABE, 2013; BORANGA; SALOMÃO, 2009). Contudo é de se observar que os 229 códigos para detalhar a informação dos resultados podem ser um número grande demais que dificulte a classificação e fatalmente a inserção com qualidade de dados nos sistemas de informação do Judiciário.

4 CONCLUSÕES

O presente artigo demonstrou que a utilização de metadados para organizar e recuperar informações de processos judiciais é de extrema relevância. Ficou expresso que a opção de estruturar a tabela de classe processual permite a formação de lacunas de informação sobre o conteúdo dos pedidos nos processos judiciais, o que pode se verificar também no futuro com outras tabelas.

Também se identificou que a tabela que exige maior atenção é a de assuntos, devido à sua complexidade e alta variedade de informações existentes. Assim, podemos concluir que a realização de pesquisas de campo com usuários das informações dos processos pode resultar em importantes propostas de aperfeiçoamento da arquitetura de metadados.

REFERÊNCIAS

ANDRETA, C.A. Legislação como assunto: uma proposta de extensão para a Classificação Decimal de Direito. In: PASSOS, E. (Org.). **Informação Jurídica: teoria e prática**. Brasília: Thesaurus, 2004, p. 79-124

BORANGA, A. L.; SALOMÃO, S. V. (Coord.). **Administração Legal para Advogados**. São Paulo: Saraiva, 2009.

CAMPOS, L. F. B. Metadados Digitais: revisão bibliográfica da evolução e tendências por meio de categorias funcionais. **Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, n. 23, 2007.

CONSELHO NACIONAL DE JUSTIÇA. **Implantação das tabelas processuais unificadas**. Brasília: Departamento de Pesquisas Judiciárias, 2010.

CONSELHO NACIONAL DE JUSTIÇA. **Justiça em Números 2013: ano-base 2012**. Brasília: CNJ, 2013.

CONSELHO NACIONAL DE JUSTIÇA. **Modelo de Interoperabilidade de Dados do Poder Judiciário e Órgãos da Administração da Justiça**. Brasília: CNJ, 2012.

CONSELHO NACIONAL DE JUSTIÇA. **Tabelas Processuais Unificadas**. Brasília: CNJ, 2014. Disponível em: <<http://www.cnj.jus.br/sgt/versoes.php>>. Acesso em: 1 jul.2014.

CONSELHO NACIONAL DE JUSTIÇA. **Termo de Acordo de Cooperação Técnica nº 059/2009**. Brasília: CNJ, 2009.

FREIRE, A. C. L. Administração judiciária. **Revista da Escola de Magistratura Federal da 5ª Região**, n. 8, p. 53-66, 2004. Disponível em <<http://www.trf5.jus.br/downloads/rev08.pdf>>. Acesso em: 1.jul.2014.

GILLILAND, A. J. *et. ali*. **Introduction to metadata**. 2ª ed. Los Angeles: Getty Publications, 2008. Disponível em: <<http://d2aohiyo3d3idm.cloudfront.net/publications/virtuallibrary/0892368969.pdf>>. Acesso em: 1.jul.2014.

MIRANDA, P. **Comentários ao Código de processo civil**. Rio de Janeiro: Forense, 2002.

NOGUEIRA, E.G. **Sistema de Gestão de Unidade Judicial**. 2010. Dissertação – Escola de Direito FGV Direito Rio, FGV, Rio de Janeiro, 2010.

PEREIRA, S. T. Processo eletrônico: máxima automação, extraoperabilidade, imaginalização mínima e máximo apoio ao juiz. **Revista do Tribunal Regional do Trabalho da 13ª Região**, João Pessoa, v. 16, n. 1, 2009.

SILVA, P. E. A. **Gerenciamento de processos judiciais**. São Paulo: Saraiva, 2010.

SILVA, P. E. A. *et al.* (Orgs.). Litigiosidade, morosidade e litigância repetitiva no Judiciário: uma análise empírica. São Paulo : Saraiva, 2012.

VELLUCCI, S.L. Metadata. *Annual Review of Information Science and Technology*, Medford, v. 33, 1998.

WATANABE, E. A Gestão por Resultados da Procuradoria-Geral da União. **Publicações da Escola da AGU: 20 Anos da Advocacia-Geral da União: história e propostas para o futuro**, Brasília, ano V, n. 27, 2013, p. 97-126.

A CIÊNCIA DIPLOMÁTICA PARA UMA ANÁLISE DE DOCUMENTOS FOTOGRAFICOS: UMA REFLEXÃO PAUTADA NA PUBLICIDADE

DIPLOMATICS SCIENCE FOR ANALYSIS OF PHOTOGRAPHIC DOCUMENTS: A GUIDED REFLECTION WITHIN THE ADVERTISING

Aline Ribeiro de Farias
Telma Campanha de Carvalho Madio

Resumo: A ciência contemporânea Diplomática é cada vez mais aplicada a documentos fotográficos, com o objetivo de compreender o contexto documental e, especialmente, a compreensão da fotografia enquanto original. A partir daí, há um pretexto para discutir no âmbito da Ciência da Informação, com ênfase em uma organização publicitária, qual documento, negativo, fotografia ou produto final é considerado natureza arquivística.

Palavras-chave: Diplomática. Documento Fotográfico. Publicidade.

Abstract: The Diplomatic contemporary science is increasing applied to photographic documents, aiming to the understanding of documentary context and especially the understanding of the original photo. From this, there is a pretext to discuss the scope of the Information Science with emphasis on an advertising organization, which document, negative, photograph or final product is considered archival nature.

Keywords: Diplomatic. Photographic Document. Advertising.

1 INTRODUÇÃO

Estamos diante de um avanço tecnológico estrondoso, onde a busca por informações ocorrem a todo o momento. Raras as vezes nos damos conta de que as informações as quais buscamos podem não ter uma confiabilidade e nos trazem diversos problemas. Graças a essas adversidades e com a vinda da Lei de Acesso à Informação²¹², cada vez mais existe uma procura pelos arquivos e pelas práticas arquivísticas que garantam o valor legal e fidedigno aos documentos. Isso é visualizado por Duranti ao dizer que: Essa capacidade dos registros documentais de capturar os fatos, suas causas e consequências, e de preservar e estender no tempo a memória e a evidência desses fatos, deriva da relação especial entre os documentos e a atividade da qual eles resultam, relação essa que é plenamente explorada no nível teórico pela Diplomática e no nível prático por numerosas leis nacionais. (DURANTI, 1994).

Para respaldar a Arquivologia na garantia dessas informações, a mesma se vincula à Diplomática, utilizando-a como uma ferramenta científica para auxiliar na tomada de decisão

²¹² Lei regulamentadora do direito constitucional de acesso dos cidadãos às informações públicas e seus dispositivos, no que tange os três Poderes da União, Estados, Distrito Federal e Municípios. Estabelecidos pela Diplomática para encontrar a autenticidade dos documentos.

sobre as questões em torno da gênese e análise documental da autenticidade dos documentos contemporâneos.

2 CIÊNCIA DIPLOMÁTICA CONTEMPORÂNEA

Com o propósito de distinguir a autenticidade documental e analisar a questão da falsificação nos documentos medievais, a Diplomática surge em meados do século XVII. Tem por definição a ocupação “[...] da estrutura do ato formal dos atos escritos de origem governamental ou notarial [...] para efeito de validade” (BELOTTO, 2008, p. 01) e ainda podem “fornecer a descrição e a explicação da estrutura formal dos atos escritos, sua autenticidade e fidedignidade” (BELOTTO, 2008, p. 01). Tem como foco de estudo a estrutura formal do objeto da Arquivologia, a “[...] unidade arquivística elementar [...]” (CARUCCI, 1987 apud BELOTTO, 2008, p. 04) onde se definirá a espécie documental e irá se concentrar “[...] na gênese, na constituição interna na transmissão e na relação dos documentos entre seu criador e o seu próprio conteúdo, com a finalidade de identificar, avaliar e demonstrar a sua verdadeira natureza” (DURANTI, 1995 apud BELOTTO, 2008, p. 05). Voltada desde o começo ao formato textual, a análise diplomática passa por uma reconfiguração e introduz em meados da década de 1970, estudos voltados aos novos formatos documentais. Esses começaram a aparecer frequentemente nos ambientes arquivísticos, sendo necessária uma adaptação também da Arquivologia a essa nova realidade. Para que houvesse uma aplicabilidade segura, utilizou-se dos preceitos diplomáticos, mas de uma forma contemporânea, para assegurar a compreensão da gênese documental e o entendimento dos conjuntos documentais, proposta por Luciana Duranti ao relatar:

As fontes usadas para chegar à proveniência de um fundo e obter o conhecimento das funções são confiáveis, porém não suficientes. É necessário conhecer as atividades específicas de cada organismo e isso só é possível a partir das informações reveladas no próprio documento (DURANTI, 1995, p. 202).

Esse novo método foi instaurado como estudo da Tipologia Documental, onde através da análise tipológica, é possível que o profissional arquivista, entenda o processo de criação, a funcionalidade e o contexto de forma evidente através dos elementos internos e externos. Dessa nova reconfiguração surge uma estrutura crítica de análise da Diplomática nos documentos, sendo exposta da seguinte forma:

- Elementos externos (suporte, escrita, linguagem, sinais especiais, selos, anotações);
elementos internos (protocolo e subseções, texto e subseções, escatocolo e subseções);

- Pessoas (autor do ato, autor do documento, destinatário do ato, destinatário do documento, escritor, controlador de assinatura, testemunho);
- Qualificação de assinaturas (títulos e capacidade das pessoas envolvidas);
- Tipo de ato (simples, contratual, coletivo, múltiplo, complexo ou de procedimento);
- Nome do ato;
- Relação entre documento e procedimento (especificação da fase do procedimento geral sobre o que se refere o documento e, se o documento surge de um “ato de procedimento”, a fase do procedimento correspondente);
- Tipo de documento (nome – carta, escritura; natureza - público ou privado; ação – dispositivo ou probatório; estado – original, cópia)
- Descrição diplomática (contexto – ano, mês, dia, lugar; ação – pessoas, atos; documento – nome da forma, natureza, função, estado, suporte, quantidade);
- Comentários conclusivos (DURANTI *apud* TOGNOLLI, 2010, p.124).

Através dessa análise é possível a reconstrução do contexto de produção do conjunto documental, mostrando a possibilidade de estudo das etapas nos novos modelos documentais, configurados a cada realidade e sendo necessária a Arquivologia, pois:

[...] a análise das características formais e substanciais dos documentos é indispensável para identificar as séries, para reconduzir documentos singulares aos arquivos de origem, para entender em quais relações se encontram os documentos referentes a um mesmo procedimento conservados em séries distintas. (CARUCCI, 1987 *apud* TOGNOLI, 2010, p.92).

3 DOCUMENTO FOTOGRÁFICO ANALISADO

De acordo com a breve introdução, três tipos diferentes de documentos fotográficos foram encontrados, sendo eles: Negativo Fotográfico, Positivo Fotográfico e Imagem Publicitária. Primeiramente, o Dicionário de Terminologia Arquivística (2005), caracteriza negativo como sendo uma “Imagem fotográfica formada quando se impressiona diretamente uma chapa ou um filme (1), em que os tons claros e escuros do objeto aparecem invertidos”. Desse formato, podem ser reproduzidas, desde que não esteja danificado, diversas cópias (ampliações) positivas do mesmo conteúdo, divergindo apenas em questões de qualidade, coloração, etc.

O positivo fotográfico é conceituado como sendo a “Imagem fotográfica que tem a polaridade do original (1)” (ARQUIVO NACIONAL, 2005), mais precisamente, é a ampliação ou contato que apresenta as mesmas formas e tonalidades da cena retratada. Entretanto, ambos os formatos utilizam-se da imagem como matéria-prima do conteúdo.

Considerada como uma “Representação gráfica, plástica ou fotográfica de seres, objetos ou fatos” (ARQUIVO NACIONAL, 2005), ela é de grande utilidade, uma vez que é um dos produtos mais empregados pela Publicidade e Propaganda atualmente. Esses três modelos são consumidos diariamente pela sociedade, apesar disso, cada vez mais o negativo, que é resultado do processo analógico fotográfico cai em desuso para os procedimentos tecnológicos digitais. Um dos exemplos onde a fotografia ganha maior enfoque de prova é nos processos judiciais, onde a mesma deve estar segundo o Código de Processo Civil (CPC) art. 385, § 1º, in verbis: “§ 1º. Quando se tratar de fotografia, esta terá de ser acompanhada do respectivo negativo”. Dessa forma, a segurança do atestado de inexistência de alteração ou manipulação imagética é quase que garantida.

Quando em formato digital, para ganhar uma validade ainda não reconhecida, deve estar munida dos instrumentos mecânicos de produção. Conquanto, devemos reconhecer que a publicidade utiliza-se das mais variadas formas de alteração e manipulação para chegar ao gosto ideal dos consumidores e fazer com que a imagem veiculada pela propaganda do produto seja um objeto de desejo do público alvo. Costa confirma essa intenção ao analisar a intencionalidade de sensações e o destino a ser compreendido pelo receptor, não deixando de lado o que a publicidade quer apresentar: “representações sociais, a imaginação coletiva”, (COSTA, 2008, p. 89). Para discutirmos sobre Diplomática incorporada nos Documentos Fotográficos, faremos antes um exercício de análise diplomática com um dos três formatos salientados acima. Exercício de estruturação da análise proposta por Luciana Duranti:

FIGURA 1 - Negativo fotográfico



Fonte: pt.wikipedia.org

1. Elementos Externos

1.1 Suporte – Triacetato de Celulose

1.2 Linguagem – Imagética

1.3 Sinais especiais, selos e anotações – Dados sobre o filme fotográfico;

2. Elementos Internos

- 2.1 Protocolo - Início do rolo fotográfico (ex. primeiro fotograma após informações de caráter externo)
- 2.2 Texto - Conteúdo imagético representado, com especificação
- 2.3 Escatocolo – Final do filme fotográfico (ex. filme com trinta cenas, 30º fotograma);
- 3. Pessoas
 - 3.1 Autor do ato – Publicitário responsável pela propaganda
 - 3.2 Autor do documento – Fotógrafo
 - 3.3 Destinatário do Ato – Editor de Imagens
 - 3.4 Destinatário do documento – Publicitário
 - 3.5 Escritor – Fotógrafo
 - 3.6 Testemunho – Participantes da etapa fotográfica;
- 4. Qualificação de assinaturas - Fotógrafo competente para o ramo da Publicidade e Propaganda;
- 5. Tipo de ato – Depende da cena a ser retratada;
- 6. Nome do Ato – Sessão de fotografias para propaganda (...);
- 7. Relação entre documento e procedimento – fase inicial de obtenção da imagem para o produto final;
- 8. Tipo de documento
 - 8.1 Nome – Negativo
 - 8.2 Natureza – Privado
 - 8.3 Ação – Probatório do ato fotográfico
 - 8.4 Estado – Original;
- 9. Descrição diplomática
 - 9.1 Contexto - Datação da imagem (dia, mês e ano)
 - 9.2 Ação – Fotografia com a representação de algo para determinada publicidade
 - 9.3 Documento – Original
 - 9.4 Função – Oferecer a matéria prima para o produto final
 - 9.5 Quantidade – Quantidade de filmes e de fotogramas;
- 10. Comentários conclusivos – Breve descrição do conteúdo imagético

A utilização da estruturação de análise proposta por Duranti é a mesma em qualquer fase documental, a diferenciação ocorre na descrição dos metadados utilizados nessa estruturação, como suporte, utilização de diferentes linguagens junto à linguagem imagética, elementos de composição do documento e principalmente tipo documental.

Através desse exercício observa-se a possibilidade da utilização da Diplomática para com os documentos contemporâneos. Esses documentos são de arquivos, pois são:

[...] produzidos por uma entidade pública ou privada ou por uma família ou pessoa no transcurso das funções que justificam sua existência como tal,

guardando esses documentos relações orgânicas entre si. Surgem, pois, por motivos funcionais administrativos ou legais. Tratam, sobretudo de provar, testemunhar alguma coisa. Sua apresentação pode ser manuscrita, impressa ou audiovisual; são em geral exemplares únicos e sua gama é variadíssima, assim como sua forma e suporte (BELOTTO, 2004, p.37).

Esses três exemplos comprovam momentos e atividades distintas dentro de uma agência publicitária, por isso devem ganhar o reconhecimento dito por Delmas, que expõe que os documentos orgânicos (ou seja, arquivísticos), produzidos no curso de uma atividade, carregam a partir de seu contexto de gênese, sinais de validação: "o fato de não serem atos gratuitos confere à informação neles contida a presunção de autenticidade" (DELMAS, 1996 apud CAMARGO, 1998, p. 172). E na informação orgânica e estruturada, e não mais no meio de veicular, que se encontra a autenticidade de um documento. Essa autenticidade, segundo Bearman (1994 apud CAMARGO, 1998, p. 173), é quem cede a possibilidade de reconhecimento do documento, independente se o mesmo passou por processos que modificaram a veracidade de seu conteúdo quando comparado ao real, no caso da publicidade.

4 CONCLUSÕES

Em síntese, após refletir, temos por conclusão de que os três exemplos documentais são originais em determinado contexto de sua produção, pois comprovam determinadas atividades em momentos e estágios diferenciados, para a elaboração do produto final a ser veiculado aos consumidores. São também passíveis de análises diplomáticas, abordando explicitamente os laços dos componentes com as ações desenvolvidas, as relações dos tipos documentais com os tipos de procedimentos envolvidos para a produção do mesmo, nunca deixando de compreender o início dessa relação. A utilização da ciência Diplomática pela Arquivologia, também norteia, de forma consciente, o reconhecimento das espécies e tipos documentais, orientando de forma controlada o desenvolver desses especificadamente. E ainda, tendo em vista o âmbito do profissional, o mesmo não deve deixar de pensar em novos documentos, mesmo com as modificações físicas. Além do que, esses profissionais são detentores e gestores das informações, e devem disseminá-las de forma eficiente e eficaz para quem dela necessitar. Para isso a tecnologia vem favorecer o profissional, pois consegue distribuir as informações que custodia em um espaço de tempo abrangente, sempre acautelando sobre a legalidade e a intencionalidade dessas informações.

REFERÊNCIAS

ARQUIVO NACIONAL (Brasil). Dicionário brasileiro de terminologia arquivística. Rio de Janeiro: Arquivo Nacional, 2005.

BEARMAN, David. Electronic evidence: strategies for managing records in contemporary organizations. **Archives & Museum Informatics**. Pittsburgh, 1994.

BELLOTTO, Heloísa Liberalli. **Arquivos permanentes: tratamento documental**. 2. ed. Rio de Janeiro: FGV, 2004.

BELLOTTO, Heloísa Liberalli. **Diplomática e tipologia documental em arquivos**. 2. ed. Brasília, DF: Briquet de Lemos, 2008.

CAMARGO, A. M. de A. Contribuição para uma abordagem diplomática dos arquivos pessoais. **Estudos Históricos**, Rio de Janeiro, v. 11, n. 21, p. 169-174, 1998.

CARUCCI, P. **II documento contemporaneo: Diplomatica e critério di edizione**. Roma: La Nuova Italia Scientifica, 1987.

COSTA, Joan. **A imagem da marca: um fenômeno social**. São Paulo: Edições Rosari, 2008.

DELMAS, Bruno. Manifesto for a contemporary diplomatics: from institutional documents to organic information. **The American Archivist**, Chicago, v. 59, n. 4. 1996.

DURANTI, L. **Diplomatica: usos nuevos para uma antiga ciência**. Carmona, Sevilla: S&V Ediciones, 1995.

DURANTI, L. Registros documentais contemporâneos como prova de ação. **Estudos Históricos**, v. 7, n. 13, 1994.

RODRIGUES, Ana Célia. **Diplomática contemporânea como fundamento metodológico da identificação de tipologia documental em arquivos**. 2008. 258f. Tese (Doutorado em História Social) - Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2008.

TOGNOLI, N. B. **A contribuição epistemológica canadense para a construção da arquivística contemporânea**. 2010. 199f. Dissertação (Mestrado em Ciência da Informação). Universidade Estadual Paulista, Marília, 2010.

TOGNOLI, N. B. GUIMARÃES, J. A. C. A organização do conhecimento arquivístico: perspectivas de renovação a partir das abordagens científicas canadenses. **Perspectivas em Ciência da Informação**, v. 16, n. 1, p. 21-44, jan./mar. 2011.

O VOCABULÁRIO CONTROLADO COMO FERRAMENTA DO PROCESSO DE ORGANIZAÇÃO E RECUPERAÇÃO DA INFORMAÇÃO

CONTROLLED VOCABULARY AS A TOOL FOR INFORMATION ORGANIZATION AND RETRIEVAL

Luciana Davanzo
Walter Moreira

Resumo: A adoção de vocabulários controlados no processo de gestão documental tem sido tomada como condição necessária ao pleno funcionamento do sistema e à eficácia nos processos de representação e recuperação da informação. No cenário que envolve os ambientes digitais, a temática ganha novos contornos em função, principalmente, da velocidade e da profusão documental que lhes são características. Desse modo, investigou-se os modos pelos quais o vocabulário controlado pode contribuir no processo de organização e recuperação da informação. Para tanto, utilizou-se como procedimento metodológico a pesquisa exploratória, de postura reflexiva e baseada em pesquisa bibliográfica, com identificação de documentos pertinentes à temática e ao desenvolvimento dessa reflexão e com potencial para subsidiar a continuidade do estudo. A organização das proposições no discurso da pesquisa permitiu confirmar o protagonismo e a necessidade do vocabulário controlado nas ações de representação da informação, quer seja no momento da indexação realizada pelo profissional da informação, quer seja no momento da recuperação da informação pelo usuário.

Palavras-chave: Organização da informação. Recuperação da informação. Vocabulário controlado.

Abstract: The adoption of controlled vocabularies in documentary information systems has been taken as something required for the full operation of the system and its efficiency in the processes of representation and information retrieval. In the scenario that involves digital environments, this topic gains new contours due, mainly, to its characteristics of speed and profusion of documents. Thus, we investigated the ways in which the controlled vocabulary can contribute to the organization and information retrieval processes. For this, we used as a methodological procedure an exploratory research, in a reflexive way, and based on bibliographic research. In doing so, we identified the documents that were relevant to support this specific work and to the continuation of the major study. The organization of propositions in the discourse of the research allowed us to confirm the role and the need of controlled vocabularies in activities such as information representation, whether at the time of indexing performed by an information professional, whether at the time of information retrieval by the user.

Keywords: Information organization, Information retrieval, Controlled Vocabulary.

1 INTRODUÇÃO

Tornou-se lugar comum, mas é sempre útil lembrar: o uso cada vez maior dos recursos automáticos e dos ambientes digitais vem alterando exponencialmente os modos pelos quais as informações são produzidas, organizadas, disseminadas e também recuperadas.

Contudo, embora existam mecanismos automáticos que atuam também como facilitadores na interação entre os usuários e os estoques de informação, a garantia de

qualidade dessa interação ainda ocorre, basicamente, por meio de processos humanos de tratamento temático da informação, em tarefas que envolvem análise criteriosa e procedimentos de inferência e de controle de vocabulário.

Tendo-se em vista a maior ou menor complexidade dos fluxos informacionais, portanto, a organização da informação tem como objetivo "dar suporte ao fluxo de tratamento e recuperação dos objetos informacionais estruturados, semi-estruturados e não estruturados nas organizações" (VICTORINO; MEDEIROS, 2009, p.8).

Nesse trabalho, os vocabulários controlados são abordados considerando-se sua utilização como instrumentos auxiliares ao processo de organização e recuperação da informação arquivística. Considera-se também a função que desempenham no processo de gestão documental, que se refere ao controle, à padronização e à normalização dos termos (descritores e não-descritores) que irão compor o repertório terminológico a ser utilizado para fins de representação e recuperação da informação.

Os vocabulários controlados atuam como linguagem documentária e são construídos especialmente para "padronizar e facilitar a entrada e a saída de dados em um sistema de informações" (KOBASHI, 2008). Essa padronização visa a aumentar a precisão do sistema e, por extensão, sua eficácia e sua confiabilidade. Um vocabulário controlado deve possibilitar, portanto, a representação da informação e o controle dessa representação em todos os níveis de interação do sistema.

Os conceitos de "vocabulário controlado" e de "linguagem documentária" estão interligados e chegam mesmo a confundir-se, principalmente para o usuário não especializado. Uma linguagem documentária é, entretanto, por definição, mais ampla.

A linguagem documentária, conforme a conceituação de Lara (2004, p.1),

refere-se ao conjunto de diferentes tipos de instrumentos especializados no tratamento da informação bibliográfica (sistema de classificação enciclopédica ou facetada e tesouros) designa de modo mais amplo e completo, a linguagem especialmente construída para organizar e facilitar o acesso e a transferência da informação.

Os vocabulários controlados exercem papel de destaque em relação à organização dos documentos que estão nos ambientes informacionais, digitais ou analógicos, uma vez que, fornecem subsídios para que as informações sejam recuperadas de forma precisa.

A precisão na recuperação da informação torna-se necessária devido à constante produção dos documentos, não sendo mais desejável priorizar aspectos de alta revocação em detrimento da precisão nos resultados. Desse modo, como problemática desse trabalho, tem-se

o seguinte questionamento: como representar o conteúdo dos documentos arquivísticos de forma que os usuários possam ter acesso rápido e confiável a esses objetos informacionais?

Objetiva-se, portanto, com essa pesquisa, analisar a função do vocabulário controlado no processo de organização e recuperação da informação em relação aos conjuntos documentais que são produzidos diariamente e que, em um momento posterior, precisarão ser recuperados de acordo com as necessidades específicas dos usuários e com as intenções que definiram a criação do documento.

Justifica-se o desenvolvimento desse trabalho em função da necessidade dos documentos analógicos e/ou digitais receberem tratamento adequado a cada tipo de suporte, para que essas informações possam ser recuperadas, contribuindo para o acesso mais qualitativo da informação, gerando dessa maneira seu compartilhamento/difusão.

Nessa perspectiva, concorda-se com Narukawa (2011, p. 15-16), quando aponta que

A necessidade de tornar a informação disponível e, mais do que isso, acessível, e de permitir o seu uso pelas pessoas, deve considerar todas as questões que perpassam os processos de produção, coleta, tratamento ou organização, recuperação, disseminação e uso da informação, entendendo-se que, a partir desse uso, um novo conhecimento pode ser gerado, propiciando uma nova produção de informação.

Com esse intuito e para melhor esclarecimento da temática abordada, o presente trabalho está dividido da seguinte forma: na primeira parte foi tratado o assunto relacionado a organização e recuperação da informação, na qual são abordadas também questões referentes ao seu objetivo e uso; na segunda parte foram apresentadas algumas contribuições em relação ao uso do vocabulário controlado correlacionando-os com a organização e recuperação da informação.

2 ORGANIZAÇÃO E RECUPERAÇÃO DA INFORMAÇÃO

Com o aumento do volume de informação em circulação, independentemente do fato de que isso possa ocorrer em suporte analógico ou digital ou mesmo da discutível qualidade dessa informação, aumenta também o número de seus usuários, reais ou potenciais, e são requeridos processos e instrumentos que possam atuar de modo mais dinâmico na criação das condições necessárias à produção e circulação do conhecimento. Coloca-se, então, como condição do uso da informação, sua eficaz organização.

Nessa perspectiva, conforme observou Aguiar (2008) e Alvarenga (2003):

Torna-se necessário que os profissionais da informação desenvolvam e implementem sistemas representacionais que estabeleçam a confluência entre a organização cognitiva imposto ao conhecimento pelo seu produtor (representação primária) e a organização conceitual imposta ao documento pelo especialista da informação (representação secundária). A confluência

entre essas duas estruturas de representação constitui-se no componente central do trabalho de arquivistas, bibliotecários e demais profissionais da informação [...] (ALVARENGA, 2003, p. 12).

Não se admite como produtivo, no universo da ciência da informação, a existência de ambientes documentários, analógicos ou digitais, que não contemplem nenhum modelo de organização da informação. Tendo-se em vista que os procedimentos de organização perpassam todo o ciclo documental, ou seja, produção, registro, organização, disseminação, recuperação e compartilhamento da informação, torna-se vez mais relevante que existam instrumentos que garantam não somente a precisão na recuperação da informação, mas também a padronização da linguagem utilizada.

Neste sentido, o vocabulário controlado agrega valor ao processo de organização e recuperação da informação porque sua utilização possibilitará o controle do repertório terminológico que será utilizado para descrever e/ou representar os conceitos que descrevem um determinado campo. Deve-se ressaltar, portanto, que a adoção de uma linguagem clara, padronizada e, na medida do possível, comum a todos diminuirá consideravelmente as dificuldades de comunicação que efetivamente existem entre usuários e sistema.

Concorda-se, portanto, com Aguiar (2008, p.114), quando afirma que "os processos de análise documentária e as linguagens de organização e representação da informação são mecanismos indispensáveis para dar consistência aos processos de organização, recuperação e transferência de informação" [...].

3 O USO DO VOCABULÁRIO CONTROLADO NA ORGANIZAÇÃO DA INFORMAÇÃO

A elaboração de vocabulário controlado está fundamentada no princípio de que se trata de "um instrumento dinâmico, capaz de ser atualizado de forma criteriosa, [que] requer uma estrutura de relações lógico-semânticas explícitas entre a macroestrutura e microestrutura, em seus diferentes níveis" (OLIVER, 2009, p.6).

Do mesmo modo como auxiliam o indexador no momento de construção da representação documentária, os vocabulários controlados auxiliam o usuário (que também pode ser o indexador) no momento da recuperação da informação. A adoção de vocabulário controlado na entrada e na saída do sistema minimiza a possibilidade de que a informação existente no sistema não seja encontrada ou que sejam encontradas informações que não atendam aos interesses de busca do usuário. Através da construção dos vocabulários controlados será possível garantir uma maior e melhor comunicação entre os usuários e sistema.

Nesta perspectiva, Smit e Kobashi (2008, p.1) esclarecem que vocabulário controlado é "uma linguagem artificial constituída de termos organizados em estrutura relacional" sendo utilizado para "padronizar e facilitar a entrada e saída dos dados em um sistema de informações".

Falar em sistema implica, naturalmente, o conhecimento das variáveis que lhe afetam. Por esse motivo, os vocabulários controlados devem ser elaborados de acordo com as políticas, estrutura, missão e valores de cada instituição, observando-se também seus usuários e seus conjuntos documentais.

Diante do exposto, pode-se inferir que, a padronização e, por consequência, o controle da linguagem utilizada nos vocabulários controlados contribuem também para eliminar ou evitar a dispersão de informações ao mesmo tempo em que contribui para que os usuários tenham confiança no sistema, conforme salienta Oliver (2009).

A falta de padronização relacionada às terminologia científica expressa nos documentos também corroboram para que existam problemas relacionados a recuperação da informação. Portanto, "representar o conteúdo informacional de um determinado documento com características que demonstram a facilidade do acesso é essencialmente importante" (ROCHA; DOMINGUES; CERVANTES, 2013, p.475).

A partir da modernização dos instrumentos que auxiliam no processo de organização e recuperação da informação torna-se possível garantir o acesso aos conjuntos informacionais. Portanto, aprimorar os instrumentos de representação e recuperação da informação, mais do que necessário é imperativo.

Os vocabulários controlados apresentam-se como instrumentos eficazes para aplicação no processo de representação da informação, assim como em relação ao processo de recuperação da informação, uma vez que proporcionam meios para controlar a parametrização das linguagens (do documento, do usuário e do sistema) que serão empregadas.

4 CONSIDERAÇÕES FINAIS

Notadamente na área de arquivologia, ainda existem muitas lacunas no que se refere ao desenvolvimento e à aplicação de vocabulários controlados. Em comparação com a quantidade de estudos desenvolvidos na biblioteconomia, há um número ainda muito pequeno de artigos, trabalhos e acadêmicos e trabalhos apresentados em congresso que abordam os aspectos teóricos e metodológicos, de desenvolvimento e aplicação de vocabulários controlados em arquivos. Isso está relacionado, sem dúvida, também ao processo de consolidação das pesquisas em nível de pós-graduação desenvolvidas em programas de

mestrado e doutorado, cujo foco seja as questões arquivísticas e os recortes temáticos que caracterizam a área.

Em trabalhos futuros pretende-se inventariar e sistematizar a produção científica referente às relações entre vocabulário controlado e arquivologia, observando-se a produção brasileira e de outros países, em busca da construção de um referencial teórico atualizado e consistente e do mapeamento dos núcleos de excelência em pesquisas sobre o assunto.

Dessa forma, acredita-se que o trabalho apresentado – originado de pesquisa exploratória e tomado na condição de subsídio para o desenvolvimento de pesquisa mais ampla que ora se realiza em nível de pós-graduação (mestrado) – contribuiu para a consecução do debate sobre a utilização de vocabulários controlados na construção de representações da informação, seja na fase de organização da informação, seja na fase de sua recuperação.

REFERÊNCIAS

AGUIAR, F. L. **O controle de vocabulário como dispositivo metodológico para a organização, tratamento e recuperação da informação arquivística**. Dissertação (Mestrado em Ciência da Informação) - Pontifícia Universidade Católica de Campinas, Campinas, 2008.

ALVARENGA, L. Representação do conhecimento na perspectiva da ciência da informação em tempo e espaço digitais. **Encontros Bibli**, n.15, p.1-23, jan./jun. 2003.

KOBASHI, N. Y. **Vocabulário controlado: estrutura e utilização**. Brasília, Escola Nacional de Administração Pública, 2008. Disponível em: <http://www2.ena.gov.br/rede_escolas/arquivos/vocabulario_controlado.pdf>. Acesso em: 09 set. 2014.

LARA, M. L. G. Linguagem documentária e terminologia. **Transinformação**, v.16, n.3, p.231-240, set./dez., 2004.

NARUKAWA, C. M. **Estudo de vocabulário controlado na indexação automática: aplicação o processo de indexação do sistema de indexação semiautomática (SISA)**. Dissertação (Mestrado em Ciência da Informação) – Universidade Estadual Paulista, Marília, 2011.

OLIVER, P. R. C. Vocabulário controlado. **ECM Connection**, 2009. Disponível em: <<http://ecmconnection.ning.com/group/classificaotaxonomiaematadados/forum/topics/vocabulario-controlado>>. Acesso em: 10 jun. 2014.

ROCHA, J. R.; DOMINGUES, G.; CERVANTES, B. M. N. Estudo de metodologias para a construção de vocabulários controlados no âmbito da Ciência da Informação. In: SEMINÁRIO EM CIÊNCIA DA INFORMAÇÃO, 5. Londrina, PR, 2013. **Anais...** Londrina, PT: UEL, 2013. p. 460-478. Disponível em: <<http://www.uel.br/eventos/cinf/index.php/secin2013/secin2013/paper/view/150/91>>. Acesso em: 26 maio 2014.

SMIT, J.W; KOBASHI, N.Y. **Como elaborar vocabulário controlado para aplicação em arquivos**. São Paulo: Arquivo do Estado, 2008. Disponível em: <http://www.arqsp.org.br/arquivos/oficinas_colecao_como_fazer/cf10.pdf>. Acesso em: 10 jul. 2014.

VICTORINO, M.C.; MEDEIROS, M.B.B. Modelagem da informação em sistemas de informações computadorizados para o reuso da informação nas organizações. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 10., 2009, João Pessoa. **Comunicações orais...** João Pessoa, 2009. p. 1-20. Disponível em: <<http://dci2.ccsa.ufpb.br:8080/jspui/handle/123456789/423>>. Acesso em: 15 abr. 2014.

TESAURO AFRO-BRASILEIRO: USO ESTRATÉGICO PARA ORGANIZAÇÃO E RECUPERAÇÃO DE INFORMAÇÃO

AFRO-BRAZILIAN TESAURO: STRATEGIC ORGANIZATION AND USE FOR INFORMATION RETRIEVAL

Graziela dos Santos Lima
Marcia Silveira Kroeff
Divino Ignacio Ribeiro Junior

Resumo: O Tesauro é uma linguagem documentária formada por descritores semanticamente relacionados que tem sido utilizado em diversos domínios do saber. Ao buscar-se tesouros que contemplam a área de estudos africanos e afro-brasileiros para auxiliar a recuperação da informação, percebeu-se que os únicos resultados voltam-se a Tesouros de Folclore e Cultura Popular Brasileira, não abrangendo muitos dos termos utilizados pelos pesquisadores no que se refere à cultura e história africana e afro-brasileira. Trata-se de estudo em desenvolvimento no Programa de Pós-Graduação em Gestão de Informação da Universidade do Estado de Santa Catarina tem por objeto identificar e organizar termos/conceitos da temática dos estudos africanos e afro-brasileiros para elaboração de um Tesauro que auxilie os bibliotecários e outros profissionais na indexação e recuperação da informação. A pesquisa tem enfoque bibliográfico, exploratório, qualitativo e quantitativo e deve ser concluída em 2016. Pretende-se construir um tesauro que integre as bases nas unidades de informação que contemplam a temática para fins de estudos e que seja relevante aos profissionais da área da informação, assim como aos pesquisadores da área da cultura e história africana e afro-brasileira.

Palavras-chave: Tesauro. Termos. Afro-Brasileiros.

Abstract: The thesaurus is a documentary language formed by semantically related descriptors that have been used in various fields of knowledge. To fetch up thesauri that include the area of African studies and african-Brazilian to help the retrieval of information, it was realized that the only return results to Tesouros Folklore and Popular Culture Brazilian does not cover many of the terms used by researchers regarding culture and African history and african-Brazilian. It is a study in development at the Graduate Program in Information Management at the University of the State of Santa Catarina's purpose is to identify and organize terms / concepts of the theme of African african-Brazilian studies and for the preparation of a thesaurus that will help them librarians and others in indexing and information retrieval. The research literature is exploratory, qualitative and quantitative approach and should be completed in 2016. Was intended to build a thesaurus that covers the bases in information units that include the topic for study purposes and which is relevant to professionals in the field of information as well as to researchers in the field of culture and African history and african-Brazilian.

Keywords: Thesaurus. Terms. African-Brazilian

1 INTRODUÇÃO

A presente pesquisa em desenvolvimento no Programa de Pós-Graduação em Gestão da Informação da Universidade do Estado de Santa Catarina tem por objeto identificar e organizar termos/conceitos da temática dos estudos africanos e afro-brasileiros para elaboração de um Tesauro que auxilie os bibliotecários e outros profissionais na indexação e recuperação da informação.

A preocupação em realizar este estudo surgiu quando, em uma pesquisa na web para identificar bibliografias na área em estudo, ao recuperar as informações, os resultados indicaram assuntos relevantes e irrelevantes, não satisfazendo adequadamente as necessidades de informação dos usuários.

Durante dois anos sendo assistente administrativa e financeira da Associação Brasileira de Pesquisadores Negros (ABPN), organismo nacional que congrega cerca de noventa e seis núcleos de estudos afro-brasileiros das mais importantes instituições de ensino superior do país, e mil e duzentos associados de diferentes áreas do conhecimento, percebi uma oportunidade para estimular o desenvolvimento de estudos e pesquisa sobre África e Diáspora Africana no Brasil.

Vivemos tempos de exaltação multiculturalista, de preocupação com a emergência de novos sujeitos de direito, com suas pautas, especificidades e reivindicações. A importância de se obter uma recuperação da informação relevante está na quantidade de documentação disponível após a explosão bibliográfica, a partir de 1980, e o tempo limitado dos pesquisadores para buscar e assimilar informação (GUEDES, 1994 apud LAPA; SANTOS e VIEIRA, 2010).

Muitas bibliotecas estão vinculadas a Núcleos de Estudos Afro-Brasileiros (NEABs) e grupos correlatos e muitas das publicações são oriundas de estudos e pesquisas em arquivos retirados de documentos dos séculos XVI a XIX. Os termos e os conceitos utilizados nesses séculos muitas vezes são termos desconhecidos pelos profissionais bibliotecários, em especial, os que lidam com a indexação, além de ser um produto inédito focado na temática em questão, o tesouro evitará erros relacionados com conceitos iguais, porém com significados correspondentes a uma época e evitará ambiguidade.

Um estudo detalhado para a construção de tesouros para estudos africanos e afro-brasileiros servirá como mecanismo de controle terminológico para facilitar a recuperação da informação em um Sistema de Recuperação de Informação, viabilizando a temática nos diversos suportes em uma unidade de informação. Desse modo, a análise dos termos para construção de um tesouro servirá para dar visibilidade à produção e ao mesmo tempo estimulará a multiplicação de estudos acerca da temática.

Este estudo, além de contribuir com a Ciência da Informação, na construção de um tesouro para uma área do conhecimento, facilitará na resolução de eventuais dúvidas com relação às palavras-chaves nas atividades de indexação em unidade de informação especializada e nas que contemplam a temática.

Por fim, essa pesquisa também objetiva contribuir como mais uma ação dentre muitas ações de caráter afirmativo desencadeadas pelo Estado nas últimas décadas, na área de Gestão de Informação. Buscamos contribuir, assim, para que a Ciência da Informação inscreva-se, de forma cada vez mais presente e completa, entre as áreas de conhecimento capazes de contribuir para a busca de uma sociedade mais igualitária.

O Tesouro é uma linguagem documentária formada por descritores semanticamente relacionados e tem sido utilizado em diversos domínios, tais como, ciência da informação, mecânica, folclore, jurídico, gênero e mulheres, dentre outras áreas do conhecimento. Além disso, é um instrumento de controle terminológico e sua estrutura possibilita processo de tradução. Vera Viana dos Santos afirma que “o princípio norteador desse trabalho traz a ideia de que os tesouros são construídos com função de organizar e representar o conhecimento registrado em áreas especializadas.” (SANTOS, 2011, p. 03).

Ao buscar-se tesouros, aqui entendidos por Sales e Café (2009, p. 232), como “vocabulário controlado formados por termos-descritores semanticamente relacionados e que atuam como instrumento de controle terminológico”, que contemplam a área de estudos africanos e afro-brasileiros para auxiliar a recuperação da informação, percebeu-se que os únicos resultados voltam-se a Tesouros de Folclore e Cultura Popular Brasileira, não abrangendo muitos dos termos utilizados pelos pesquisadores no que se refere à cultura e história africana e afro-brasileira. Curiosamente, apesar das tantas fontes existentes ainda não vislumbramos tesouros que facilitem a indexação e a recuperação das informações nas unidades de informação no que tange a temática em questão.

Em virtude do exposto, procura-se elaborar um tesouro afro-brasileiros que facilite a recuperação da informação e, ao mesmo tempo, auxilie os indexadores a selecionar termos adequados para documentos com essa temática.

2 MÉTODOS DE ORGANIZAÇÃO DO CONHECIMENTO

Na era da informação, diversas áreas do conhecimento têm se preocupado com o modo de recuperar as informações, disponibilizadas de forma dispersas, seja em rede ou em base de dados, além de distribuídas em diferentes suportes. As informações quando recuperadas são carregadas de informações relevantes e irrelevantes, dificultando a busca e o uso das informações por parte do usuário.

O vocabulário estruturado tem facilitado áreas específicas do conhecimento para a indexação e recuperação da informação. Segundo Jesus (2002, p. 5),

o uso de um vocabulário estruturado permite o pesquisador recuperar a informação com o termo exato utilizado para descrever o conteúdo daquele

documento científico, provenientes de termos consistentes, que permite ao usuário selecionar a informação de que necessita.

Dentro do vocabulário estão às linguagens documentárias que auxiliam na comunicação entre indexador, sistema e usuário, “são instrumentos de controle terminológico utilizados no processo de indexação e recuperação da informação” segundo Sales (2008, p. 16).

Das linguagens documentárias existentes, uma das mais comentadas são os tesouros, as ontologias e o sistema da classificação facetada, consistindo em “modelos de representação do conhecimento que, servindo como instrumentos de controle terminológico, auxiliam o processo de indexação e recuperação de documentos por assunto” (SALES; CAFÉ, 2008, p. 232).

Com a necessidade de recuperação da informação, os tesouros têm como característica específica estabelecer relações entre os termos que o compõem, ou seja, termos relacionados semanticamente.

por meio dos conceitos, termos e relações semânticas que as linguagens documentárias adquirem o status de sistemas representativos do conhecimento armazenado no mundo das idéias. Estes sistemas são aplicados na organização da informação com vistas a aumentar a qualidade da recuperação da informação. Desta forma, funcionam como instrumentos de comunicação e contribuem para o processo de socialização da informação e do conhecimento (GHENO, 2013, p. 14).

As ontologias, por meio de uma estrutura conceitual, também têm o mesmo objetivo do tesouro, representar e recuperar informação, porém viabilizando um acordo no uso do vocabulário compartilhado de uma maneira coerente e consistente (SALES; CAFÉ, 2008). Segundo Sales, Campos e Gomes (2008) as ontologias são mais que uma linguagem documentária devido à funcionalidade que permitem que a máquina possa processar o raciocínio automatizado, por meio de regras e inferências.

A Teoria de Classificação facetada, desenvolvida por Ranganathan em 1920, é um método de distribuição por categorização ou facetação de uma determinada área do conhecimento (NOVAES, 2009). Considerando a necessidade de obter uma recuperação da informação de forma precisa utilizando os termos da temática africana e afro-brasileira, usaremos o Tesouro como método de organização do conhecimento.

2.1 Tesouro

O primeiro registro do Tesouro ocorreu em Londres em 1852 no dicionário de Peter Mark Roget, intitulado “Thesaurus of English words and phrases” (JESUS, 2002).

Porém, este dicionário era agrupado de acordo com as ideias que os termos poderiam expressar (COELHO, et al, 2004).

Na década de 1970, começaram a empregar princípios para a elaboração de tesouros possuindo duas linhas condutoras, conceitos e usos de categorias (COELHO et al, 2004).

Carlan (2010) informa que o tesouro tem uma estrutura sintética simplificada formando cruzamento lógico e hierárquico dos descritores. Jesus (2004, p.4) afirma que o tesouro “possui uma relação associativa e sua estrutura não se baseia em conceito de palavras simplesmente, mas sim de termos conceituado e relacionados”.

No Brasil o termo tesouro empregado por Cavalcanti tem como principal definição,

Lista estruturada de termos associados, empregadas por analista de informação e indexadores, para descrever um documento com a desejada especificidade, em nível de entrada, e para permitir aos pesquisadores a recuperação da informação que procura (CAVALCANTI, 1978, p.27).

Na representação e organização do conhecimento, o tesouro tem auxiliado no controle terminológico e inserido, por conta dessa estrutura, nos sistemas de recuperação de informação das unidades de informação. Por conta disso tem o objetivo de “dar assistência ao usuário (pesquisador ou indexador) de maneira que ele consiga encontrar o termo que represente um determinado significado para o que se procura” (JESUS, 2002, p. 16).

Portanto, o estudo e aplicação de Tesouro na área africana e afro-brasileira pode ser de grande utilidade, proporcionando ao indexador e usuário uma recuperação da informação relevante.

2.2 Estudos Africanos e Afro-Brasileiros

Os estudos a cerca da temática africana e afro-brasileira são realizados à *priore* para dar suporte aos pesquisadores, profissionais do serviço público, principalmente professores das redes públicas de ensino em cumprimento a Lei Federal 10.639/03. História e Cultura Africana e Afro-Brasileira têm em seu conteúdo termos/conceitos muitas vezes utilizados de acordo com a época, expressando um determinado sentido.

Com esses estudos, apresentaremos uma profusão de novos conceitos e revisão de velhas noções de termos como “negro”, “afrodescendentes”, “preto” e “crioulo”, termos caros que, quando não utilizados na indexação, de acordo com a informação descrita no documento, podem dificultar a recuperação da informação. Segundo Cardoso e Rascke (2014, p. 14) “os termos necessitam de uma explicação e análise histórica, pois termos e palavras possuem

sentidos que se alteram ao longo do tempo”. Nesse sentido, surge a necessidade de verificar termos que auxiliem especialistas na área dos Estudos Africanos e afro-brasileiros, indexadores e usuários na recuperação de informações relevantes. A temática possui vários termos que, dependendo do documento ou texto, possuem significado relativo ao sujeito e ao espaço onde se encontram.

3 MATERIAIS E MÉTODOS

Visando ir ao encontro do objetivo da pesquisa, trata-se de uma pesquisa bibliográfica, exploratória, qualitativa e quantitativa, a ser realizada no Núcleo de Estudos Afro-Brasileiros da Universidade do Estado de Santa Catarina.

A Pesquisa bibliográfica, segundo Cervo e Bevilacqua (2002), é a explicação de um problema a partir de um referencial teórico, “busca conhecer e analisar as contribuições culturais ou científicas do passado existentes sobre um determinado assunto, tema ou problema”. Por meio de referencial teórico buscaremos delinear definições acerca de linguagem documentária, entendimento para compreender a representação do conhecimento em diferentes metodologias de construção; definição como Tesouro termo principal que desencadeará esta pesquisa; além dos estudos africanos e afro-brasileiros.

Para fundamentar a construção utilizar-se-á materiais resultantes de pesquisas de mestrado e doutorado, livros e artigos científicos, bem como materiais organizados a partir de outras experiências de pesquisa acadêmica, sendo analisados documentos para abstração de conceitos que remetem às informações relevantes para os/as usuários/as. Portanto, serão utilizadas fontes de informação primárias e secundárias, possibilitando maior abrangência e cobertura sobre a temática em questão.

Pesquisa exploratória, com o intuito de verificar termos, conceitos usados autores/as nas obras relacionadas com a temática africana e afro-brasileira, conforme o autor supracitado, a “pesquisa exploratória realiza descrições precisas da situação e quer descobrir as relações existentes entre os elementos componentes da mesma.” (CERVO; BEVIAN, 2002, p. 69).

A pesquisa qualitativa compreende, “um conjunto de diferentes técnicas interpretativas que visam a descrever e a decodificar componentes de um sistema complexo de significados” (NEVES, 1996, p. 02). A pesquisa quantitativa é “projetada para gerar medidas precisas e confiáveis que permitam uma análise estatística” (MORESI, 2003, p. 64).

Após a leitura e coleta dos termos, os mesmos serão inseridos em um software específico para construção de Tesouros. A pesquisa, no momento está na fase de revisão de literatura para a fundamentação teórica e que deve ser concluída em dois anos.

4 CONSIDERAÇÕES FINAIS

O interesse na presente pesquisa, esboçado nestas breves linhas, constitui em apresentar um instrumento que facilite na recuperação exata da informação evitando ambiguidades, além de visibilizar a temática em questão e gerar conhecimentos acerca destes estudos.

Como a área de estudos africanos e afro-brasileiros é ampla e abrange diversas regiões, pretende-se construir um tesouro que integre as bases nas unidades de informação que contemplam a temática para fins de estudos e que seja relevante aos profissionais da área da informação, assim como aos pesquisadores da área da cultura e história africana e afro-brasileira.

REFERÊNCIA

- CARDOSO, Paulino de Jesus Fracisco; RASCKE, Karla Leandro. Lei Federal 10.639/03, discussão de conceitos: multiculturalismo, diversidade, ações afirmativas, racismo, preconceito, afrodescendente, negro, entre outros. In: CARDOSO, Paulino de Jesus Fracisco; RASCKE, Karla Leandro (Org.). **Formação de professores: produção e difusão de conteúdos sobre história e cultura afro-brasileira e africana**. Casa Aberta: Itajaí, 2014.
- CARLAN, Eliana. **Sistema de organização do conhecimento: uma reflexão no contexto da ciência da informação**. 2010, 195f. Dissertação (Mestrado em Ciência da Informação), Universidade de Brasília. Pós-Graduação em Ciência da Informação, Brasília, 2010. Disponível em: <repositorio.unb.br/bitstream/10482/7465/1/2010_ElianaCarlan.pdf>. Acesso em: 2 de fev. 2014.
- CAVALCANTI, Cordelia R. **Indexação & tesouro: metodologia e técnicas**. Brasília: Associação de Bibliotecários do Distrito Federal, 1978.
- CERVO, Amado Luiz; BERVIAN, Pedro Alcino. **Metodologia científica**. 5. ed. São Paulo: Person Prentice Hall, 2002.
- GHENO, Tatiane Cristina. **Estrutura de Tesouro em Ciência da Informação: análise de tesouros nas bases de dados Lisa e Lista**. 2013, 60f. TCC (Monografia de Graduação em Biblioteconomia) – Universidade Federal de Santa Catarina. Curso de Graduação em Biblioteconomia, Florianópolis, 2013. Disponível em: <<https://repositorio.ufsc.br/handle/123456789/103711>>. Acesso em: 4 fev.2014.
- INSTITUTO DO PATRIMONIO HISTÓRICO E ARTÍSTICO NACIONAL (Brasil). Centro Nacional de Folclore e Cultura Popular. **Tesouro de folclore e cultura popular brasileira**. Rio de Janeiro, 2004. Disponível em: <<http://www.museudofolclore.com.br>> Acesso em: 2 de fev. 2014.

MORESI, Eduardo. **Metodologia de Pesquisa**. Universidade Católica de Brasília – UCB – DF, 2003. Disponível em <www.inf.ufes.br/~falbo/files/MetodologiaPesquisa-Moresi2003.pdf>. Acesso em: 22 de fev.2014.

NEVES, José Luis. **Pesquisa qualitativa: características usos e possibilidades**. Caderno de pesquisa em Administração, São Paulo, v. 1, n. 3, 2º Sem. 1996. Disponível em: <<http://www.ead.fea.usp.br/cad-pesq/arquivos/c03-art06.pdf>>. Acesso em: 26 de fev. 2014.

NOVAES, Denise Queiroz. **A Hierarquia Semântica de Conceito de Tesouros**. 2009, 146f. Tese (Doutorado em Letras) - Universidade Católica de Minas Gerais. Programa de Pós-Graduação em Letras, Belo Horizonte, 2009. Disponível em: <www.biblioteca.pucminas.br/teses/Letras_NovaesDQ_1.pdf>. Acesso em: 4 fev.2014.

SALES, Luana Farias; CAMPOS, Maria Luiza de Almeida; GOMES, Hagar Espanha. Ontologia de domínio: um estudo das relações conceituais. **Perspectivas em Ciência da informação**, v. 13, n. 2, p. 62-76, maio/ago. 2008.

SALES, R. ; CAFE, L. M. A. Os tesouros e as ontologias sob um olhar terminológico. In: Encontro Nacional de Pesquisa em Ciência da Informação - ENANCIB, 9., 2008, São Paulo. **Anais...** Diversidade Cultural e Políticas de Informação, 2008.

SALES, R.; CAFE, L. O que há de comum entre tesouros e ontologias. **Cibertextualidades**, (Porto), v. 3, p. 231-241, 2009.

SANTOS, Vera Viana dos. Construção de um tesouro especializado. In: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA, DOCUMENTAÇÃO E CIÊNCIA DA INFORMAÇÃO SISTEMAS DE INFORMAÇÃO, MULTICULTURALIDADE E INCLUSÃO SOCIAL. 2011, Alagoas. **Anais...** Maceió, Alagoas, 2011. p. 1-9.

ANÁLISE DOCUMENTÁRIA DE IMAGENS FOTOGRÁFICAS DO ARQUIVO AFONSO PEREIRA

DOCUMENTARY ANALYSIS OF PHOTOGRAPHIC IMAGES OF THE AFONSO PEREIRA ARCHIVE

Eveline Filgueiras Gonçalves
Maria Elizabeth Baltar Carneiro de Albuquerque

Resumo: Propõe uma análise das informações imagéticas do acervo fotográfico do arquivo privado pessoal do paraibano Afonso Pereira da Silva (APP), especificamente das fotografias pertinentes ao seu vínculo com a Universidade Federal da Paraíba (UFPB), e a partir delas apresentar a origem da instituição. Como objetivo geral, analisa a dimensão educacional de Afonso Pereira da Silva, por meio da sua fotoautobiografia, enfocando sua ação como educador, notadamente como partícipe da fundação da UFPB. Para alcançarmos nosso objetivo, a pesquisa pretende descrever os procedimentos de tratamento das informações imagéticas constantes do Fundo Arquivístico Afonso Pereira da Silva, bem como mapear significados e sentidos da amostra fotográfica, escolhida intencionalmente e representada por meios da grade de análise de imagens. Propomos uma pesquisa exploratório-descritiva, documental, com abordagem qualitativa. Utilizando a representação de imagens e por meio de uma imagem selecionada, trabalha-se a fotografia como artefato indiciário para ressignificação do passado, observando o referente na determinação da indiciabilidade, onde a indexação observa as regras próprias da representação através da adoção da grade de análise proposta por Barros. Acredita-se que a aplicação da proposta no acervo fotográfico do AAP tornará, no futuro, acessível a fonte de informação memorialística, por entendermos que a representação possibilita acesso aos itens documentais, permitindo novos usos e construção de conhecimentos originais.

Palavras-chave: Análise documentária de fotografias. Representação da informação imagética. Fotografia.

Abstract: It proposes an analysis of the imagetive information of the personal and private photographic archive of Afonso Pereira da Silva (APP), from Paraíba, specifically the photographs related to his bond with the Federal University of Paraíba (UFPB) and from this archive to present the institution's origin. As to the general objective, it analyzes the educational dimension of Afonso Pereira da Silva by means of his photoautobiography, focusing on his action as an educator, particularly as foundation participant of the UFPB. In order to attain our goal, the piece of research intends to describe the treatment procedures of the imagetive information present in the Afonso Pereira da Silva Archival Fund as well as to outline meanings and senses of the photographic sample chosen intentionally and represented by means of the framework of the images' analysis. An exploratory-descriptive and documental research with qualitative approach is proposed. By using the image representation and by means of a selected image, it deals with the photograph as indiciary artifact for resignifying the past, observing the referent in the index determination in which indexation perceives the rules for the representation through the adoption of the analysis framework proposed by Barros. It is believed that the use of the proposal in the APP photographic archive will enable, in the future, the access to the memorialistic information source as we understand that the representation makes possible the access to the documental items in a way to allow new uses and construction of original knowledge.

Keywords: Analysis of documentary photographs. Imagetive information representation. Photography.

1 INTRODUÇÃO

Inegavelmente vivemos numa sociedade onde a informação não possui apenas as dimensões verbais, orais e escritas, mas variados sistemas de produção de sentido, por conseguinte, nosso estar-no-mundo abarca as representações visuais dos seres animados ou inanimados, fazendo da nossa linguagem complexa, justificando que os saberes sobre linguagens não verbais não devem ser negligenciados. Assim, entendemos a relevância da nossa pesquisa visto que os estudos em representação da informação, em sua maioria, estão restritos ao universo dos documentos textuais. A representação da informação em acervos fotográficos deve ser preocupação da Ciência da Informação, pois assim será possível o compartilhamento das memórias neles contidas.

Abalizados por tais convicções, propomos uma pesquisa para a análise das fotografias do arquivo privado pessoal do paraibano Afonso Pereira da Silva. O Arquivo Afonso Pereira (AAP) está localizado na cidade de João Pessoa/PB, cujo acervo centra-se na vida e obra do deste educador, administrador, jornalista, magistrado e parlamentar.

Dessa forma, apresentamos como objetivo geral analisar a dimensão educacional do paraibano Afonso Pereira da Silva, por meio da sua fotoautobiografia feita com base nas fotografias e álbuns constantes em seu arquivo privado pessoal, enfocando sua ação de educador, especificamente como partícipe da fundação da Universidade Federal da Paraíba (UFPB). As fotografias que se encontram no referido acervo trazem, em seu conteúdo, aspectos que retratam o percurso trilhado por Afonso Pereira da Silva e é essa memória um bem precioso para a criação das identidades locais.

Para alcançarmos nosso objetivo, a pesquisa pretende descrever os procedimentos de tratamento das informações imagéticas constantes do Fundo Arquivístico Afonso Pereira da Silva, bem como mapear significados e sentidos da amostra fotográfica, escolhida intencionalmente e representada por meios da grade de análise imagens.

Se o AAP foi criado com o intuito de preservação da memória, esse deve possuir um acervo acessível para servir como subsídio para pesquisadores. Neste sentido, questionamos: é possível o acesso a fontes e conteúdos informacionais do acervo fotográfico do AAP, podendo seus itens ser identificados e localizados? Como se constitui o processo de tratamento das informações imagéticas no Fundo Arquivístico Afonso Pereira da Silva?

2 ANÁLISE DOCUMENTÁRIA DE IMAGENS

É da teoria da análise iconológica, de autoria do crítico e historiador da arte alemão Erwin Panofsky (2007), o embasamento para as diferentes metodologias apresentadas para

análise documentária de imagens. Nomeou três níveis de análise: pré-iconográfico, iconográfico e iconológico, onde a leitura iconográfica é uma análise descritiva da imagem e a iconológica uma interpretação intrínseca ao conteúdo.

Ao refletir sobre análise documentária de imagens, Smit (1987) decompôs o processo em quatro atributos: pseudotransparência, transcodificação, informações técnicas e interpretação. O primeiro atenta para o fato de a fotografia ser uma ilusão, pois não é a realidade, nem tão pouco uma cópia fiel do mundo e de seus acontecimentos. A transcodificação abarca o processo de tradução do código icônico para o verbal, apresentando as dificuldades próprias do processo: seleção de informação, perda de precisão e possibilidade de erro. A terceira propriedade adverte da importância da descrição das informações técnicas, ou seja, das condições de produção da imagem, por meio das quais ocorrem bastantes transformações. No quarto e último atributo, lembra que a interpretação demanda treino, pois é difícil e problemática a distinção entre descrição e interpretação, uma vez que a descrição, por meio da tradução cria condições para sua interpretação. Porém, por mais complexa que seja a interpretação, por incorporar valores socioculturais e ideológicos, a leitura iconológica não deve ser desprezada. Assim, fundamentada e adotando as questões elaboradas por Bléry (1976) *apud* Smit (1987, 1996), a interpretação constituiu as categorias para futura elaboração da sua grade de análise: quem, onde, quando, como e o quê.

Shatford (1984), fundamentada em Panofsky, alega que a imagem é respectivamente específica e genérica. Logo, deverá ser representada nos níveis pré-iconográfico e iconográfico, o que é diferente do significado iconológico (símbolo), que equivale ao significado expressivo, indo além da indexação baseada em conteúdo (cor, forma e textura); assim, adentra a linha de raciocínio para representação da imagem.

Partindo destes parâmetros para a identificação do conteúdo informacional da imagem, toda imagem pode ser: genericamente e especificamente **DE** alguma coisa, e **SOBRE** alguma coisa; cujo **DE** aproxima-se do referente, portanto relaciona-se aos níveis pré-iconográfico e iconográfico, e o aspecto **SOBRE** relaciona-se à interpretação do tema, logo, corresponde ao nível iconológico.


Smit (1996) apresentou proposta de grade de análise, ressaltando o imperativo da adoção de critérios para a seleção de regras para o processo de representação, objetivando alcançar consistência e eficácia no processo, renunciando estudo futuro que contemple a expressão fotográfica da imagem, ou seja, a informação imagética expressa pela técnica adotada do **COMO** foi executada.

2.1 Expressão fotográfica ou dimensão expressiva da imagem

Encontramos apontamentos sobre a importância do registro das informações técnicas da fotografia manifestada em Lacerda (1993), Smit (1997) e Manini (2002, 2004 e 2007) por oferecer melhorias para indexação de imagens, visto ampliar o conjunto de dados sobre o item analisado.

Manini (2002) oferece proposição de grade para análise documentária de fotografia, acrescentando a técnica, originando a **Dimensão Expressiva**, cuja coluna será completada com dados do quadro de recursos técnicos e suas variáveis:

QUADRO 1 – Grade de análise com dimensão expressiva

Categoria	Conteúdo Informacional		Dimensão Expressiva	Imagem	
	DE	SOBRE			
	Genérico	Específico			
<i>Quem/ O quê</i>	Homem jovem. Medalha. Evento esportivo.	Atleta (brasileiro), ginasta (brasileiro), Arthur Zanetti, medalha de ouro.	Esporte. Jogos olímpicos. Olimpíadas. Olimpíadas de Londres. Ginástica artística. Medalhista de ouro. Vitória.	Retrato. Pose. Plano fechado com detalhe. Sem profundidade.	
<i>Onde</i>		Londres, Inglaterra.			
<i>Quando</i>		2012			
<i>Como</i>	Segurando, mostrando, rindo				

Fonte: Manini (2002, p.105); Manini (2004, não paginado); Manini (2007, não paginado).

Examinando as metodologias desenvolvidas com o objetivo de estudar as informações imagéticas contidas no conjunto de fotografias pertencentes ao arquivo privado pessoal de José Simeão Leal e, com base nelas, refazer as redes de sociabilidade nas quais esse editor estava inserido, Barros construiu a seguinte assertiva em relação à proposta de Manini (2002, 2004, 2007):

Acreditamos ser relevante essa proposta na medida em que o acervo que será analisado constitui-se de fotografias produzidos através de equipamentos não digitais (o status da fotografia digital requer uma discussão mais longa, o que fugiria dos objetivos propostos), além do mais, da mesma forma que a imagem de um líder pode ser construída a partir do ângulo utilizado pelo fotógrafo (de baixo para cima, de cima para baixo, em close, através de uma lente grande-angular etc.), a forma com o indivíduo é representação [sic] através da fotografia interferem na maneira como percebemos sua inserção ou não nos ambientes e seu pertencimento aos grupos sociais dos quais supostamente foi integrante. (BARROS, 2012, p. 25).

Assim fundamentada, Barros (2012) oferece nova proposta, ampliada, de grade de análise, onde incluem categorias informacionais, elencadas por Lacerda (1993) como essenciais para analisar as informações imagéticas dos arquivos pessoais, visto que a fotografia não se limita à imagem, mas se configura também como artefato do arquivo.



Por meio da representação fotográfica, elaborando uma resignificação do fundador de instituições de ensino e educação Afonso Pereira da Silva, buscaremos revelar uma realidade ausente. As imagens que irão compor nossa amostra serão aquelas que representam um fragmento de sua vida, território específico, cuja reunião texto-visual se dará no âmbito do seu arquivo pessoal, baú de seus guardados, lugar de memória que acreditamos ser capaz de evocar a história da criação da UFPB.

3 METODOLOGIA

Propomos uma pesquisa exploratório-descritiva, documental, com abordagem qualitativa, com um *corpus* de 219 fotografias provenientes de 3 álbuns, além de 22 caixas-arquivo de documentos relacionados ao exercício profissional do pesquisado e, 36 volumes encadernados de recortes de jornais, todos referentes à UFPB.

Selecionamos uma imagem, para este pôster, do álbum 02, para realizarmos a análise documentária com fins de extrair dados sobre a origem e criação da UFPB, como apresentamos a seguir:

QUADRO 2 – Grade de análise documentária de imagem adotada

	Conteúdo Informacional	AAP alb02ft03
	<p>Conteúdo genérico/ sentido denotativo: Grupo de homens; homens sentados; traje formal; paletó; gravata; ambiente fechado; sala de reunião.</p> <p>Conteúdo específico: Professores reunidos na Faculdade de Ciências Econômicas da Paraíba.</p> <p>Onde: João Pessoa, [Faculdade de Ciências Econômicas?]</p> <p>Quando: [20/05/1952?]</p> <p>Conteúdo abstrato/ sentido conotativo: Reunião.</p>	<p>Seqüência narrativa: imagem isolada.</p> <p>Identificação dos indivíduos: Afonso Pereira da Silva</p> <p>Configuração: Participação na Reunião Ordinária da Faculdade de Ciências Econômicas da Paraíba, em que professores elegeram o Conselho Técnico da Faculdade. Nesta reunião surgiu a ideia de encaminhar uma solicitação ao Governo do Estado para oficializar a Faculdade de Direito ou a de Medicina. Com um conjunto de Universidades Estaduais, pleiteariam a federalização.</p>
<p>Autor: [s.n.]</p> <p>Informações sobre o suporte: 1 cópia, com borda branca, 21,5 x 18 cm, bom estado de conservação.</p>	<p>Dimensão expressiva: Retrato, instantâneo, luz artificial, plano fechado frontal, com profundidade de campo, P&B.</p> <p>Observações: Informações sobre a reunião estão em um recorte de jornal presente no álbum, porém sem referência, não foi possível identificar o jornal, constando o registro em caneta da data 21/05/1952.</p>	

Fonte: Barros, Azevedo Netto (2014, p. 203, adaptação nossa).

A fotografia analisada, disposta na terceira folha do álbum 02 (AAP alb02ft03), referencia a reunião ocorrida no ano de 1952, na Faculdade de Ciências Econômicas da Paraíba (1947), primeira escola de ensino superior da capital paraibana, onde surgiu a ideia de transformar as faculdades e escolas autônomas da Paraíba em estaduais. A conjectura fundamentou-se na flexibilização do governo para a criação de universidades, onde mediante o novo critério constitucional (1946), por meio da existência de uma Faculdade de Filosofia e outros dois institutos que se habilitavam para serem oficializadas como universidades. Alcançado este patamar, posteriormente solicitariam sua federalização, tornando-as Universidade Federal da Paraíba (UFPB).

5 CONSIDERAÇÕES PARCIAIS

A análise documentária, por meio da representação da informação imagética, possibilita trabalhar o artefato fotografia, como signo indiciário, revelando um passado que está representado e simbolizado na imagem.

A representação da informação imagética da fotografia selecionada nos possibilitou distinguir elementos que chegaram até nós pela percepção e fez os dados se transformarem em informações imagéticas ao definirmos respostas para satisfazer questões como: quem, onde, quando, como e o quê, sem perder de vista os aspectos baseados no **DE** e **SOBRE** o quê da imagem, ressignificando as marcas originárias da Universidade Federal da Paraíba (UFPB).

Esperamos, num futuro próximo, ampliar a aplicação da proposta no acervo fotográfico do AAP, para que esta fonte de informação memorialística fique acessível, permitindo usos e construção de novos conhecimentos, fomentando assim trabalhos na área de representação da informação e do conhecimento.

REFERÊNCIAS

BARROS, Kelly Cristiane Queiroz. **Rede humana de relações**: relações de sociabilidade a partir do acervo fotográfico de José Simeão Leal. 2012. 150f. Dissertação (Mestrado em Ciência da Informação)- Centro de Ciências Sociais Aplicadas, Universidade Federal da Paraíba, João Pessoa, 2012. Disponível em: <<http://rei.biblioteca.ufpb.br/jspui/bitstream/123456789/139/1/KCQB08032013.pdf>>. Acesso em: 21 jul. 2013.

BARROS, Kelly Cristiane Queiroz; AZEVEDO NETTO, Carlos Xavier de. O acervo fotográfico de José Simeão Leal: representação da informação imagética e descrição de redes de sociabilidade In: MOTA, Ana Roberta Sousa et. al. **Versados em ciência da informação**. João Pessoa: Imprel, 2014. cap. 9, p. 187-207.

LACERDA, Aline Lopes de. Os sentidos da imagem: fotografias em arquivos pessoais. **Acervo**: revista do Arquivo Nacional, Rio de Janeiro, v. 6, n. 1-2, p. 41-54, jan./dez. 1993. Disponível em: <http://www.arquivonacional.gov.br/media/v6_n1_2_jan_dez_1993.pdf>. Acesso em: 21 jul. 2013.

MANINI, Miriam Paula. A dimensão expressiva na indexação de documentos fotográficos. In: ENCONTRO NACIONAL DE ESTUDOS DE IMAGENS, 1., 2007, Londrina. **Anais...** Londrina: Universidade Estadual de Londrina, 2007. Disponível em: <http://repositorio.unb.br/bitstream/10482/1012/1/EVENTO_DimensaoExpressivaIndexacao.pdf>. Acesso em: 19 jul. 2013.

MANINI, Miriam Paula. Análise documentária de fotografias: leitura de imagens incluindo sua dimensão expressiva. **Cenário Arquivístico**, Brasília, DF, v. 3, n. 1, p. 16-28, 2004. Disponível em: <http://repositorio.unb.br/bitstream/10482/946/1/ARTIGO_AnaliseDocumentariaFotografia.pdf>. Acesso em: 19 jul. 2013.

MANINI, Miriam Paula. **Análise documentária de fotografias**: um referencial de leitura de imagens fotográficas para fins documentários. 2002. 226f. Tese (Doutorado em Ciências da Comunicação) - Departamento de Biblioteconomia e Documentação da Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2002. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/27/27143/tde-23032007-111516/pt-br.php>>. Acesso em: 10 jun. 2013.

PANOFSKY, Erwin. Iconografia e iconologia: uma introdução ao estudo da arte da Renascença. In: _____. **Significado nas artes visuais**. 3. ed. São Paulo: Perspectiva, 2007. cap.1.

SHATFORD, Sara. Describing a picture: a thousand words are seldom cost effective. **Cataloging & classification quarterly**, Reino Unido, v. 4, n. 4, p. 13-30, summer 1984.

SMIT, Johanna W. (Coord.) A análise da imagem: um primeiro plano. In: _____. **Análise documentária**: a análise da síntese. Brasília, DF: IBICT, 1987. p. 99-111. Disponível em: <[http://livroaberto.ibict.br/bitstream/1/1011/1/Análise documentária.pdf](http://livroaberto.ibict.br/bitstream/1/1011/1/Análise%20documentária.pdf)>. Acesso em: 15 jun. 2013.

SMIT, Johanna W. A representação da imagem. **INFORMARE**: Caderno do Programa de Pós-Graduação em Ciência da Informação, Rio de Janeiro, v. 2, n. 2, p. 28-36, jul./dez. 1996. Disponível em: <<http://ibict.phlnet.com.br/anexos/smitv2n2.pdf>>. Acesso em: 17 jun. 2013.

SMIT, Johanna W. Propositions for the treatment of iconographical information. In: CONGRESO INTERNACIONAL DE INFORMACION INFO'97, 1997, Havana. **Textos completos...** Havana: IDICT, 1997. p. 1-14.

ORGANIZAÇÃO DA INFORMAÇÃO EM CONTEXTO DINÂMICO E SENSÍVEL: TESAURO DE INTELIGÊNCIA POLICIAL

KNOWLEDGE ORGANIZATION IN A DYNAMIC AND SENSIBLE CONTEXT: THE ELECTRONIC THESAURUS OF POLICE INTELLIGENCE'S OPERATIONS (TEOIP)

Maria Aparecida Moura
Rodrigo Benjamin Gesteira

Resumo: Esse estudo, desenvolvido no âmbito da Organização da Informação, tomou como referência o caráter informacional das operações de inteligência policial e partiu da hipótese de que seria possível desenvolver uma metodologia que orientasse a elaboração de um tesouro capaz de auxiliar a atividade de inteligência policial no combate a organizações criminosas. Para tanto, observou-se as peculiaridades do contexto, no qual a informação está inserida em ambiente de diversificadas fontes de dados, revestido do sigilo e de extrema dinamicidade. Como resultado, elaborou-se o Tesouro Eletrônico de Operações de Inteligência Policial (TEOIP) que incluiu como inovação a agregação de termos oriundos da análise efetuada no linguajar da organização criminosa observada. Dessa tarefa, inferiu-se que seria possível fazer do TEOIP, além de uma ferramenta para organizar a informação, oriunda das operações de inteligência policial, um sistema de alerta do cometimento ou da preparação de determinada ação pela Orcrim. Além disso, o estudo realizado permitiu evidenciar que as práticas adotadas pela inteligência policial no combate a organizações criminosas reivindicam a composição de uma base teórica interdisciplinar e que as contribuições consolidadas no campo da Ciência da Informação, notadamente no âmbito da organização e uso da informação e do conhecimento. Dessa perspectiva, a CI pode oferecer aporte teórico-metodológico fundamental para o entendimento da ação criminosa como um fenômeno informacional dinâmico e agravado pela sensibilidade das ações que compõem a sua caracterização, monitoramento e combate.

Palavras-chave: Inteligência policial, tesouros - metodologia, informação sensível

Abstract: This study, developed in the scope of Information Organization, has taken as reference the informational side of the police intelligence's operations and started with the hypothesis that it would be possible to develop a methodology that would guide the creation of a thesaurus, capable of helping the police intelligence's activity, in fighting criminal organizations. In order to do so, the singularities of the context were analyzed, in which the information is inserted in an environment of multiple data bases, covered by secrecy and extremely dynamic. As a result, the Electronic Thesaurus of Police Intelligence's Operations (TEOIP) was developed, that includes as innovation the aggregation of terms coming from the analysis made in the language of the criminal organization observed. From this task, it was inferred that it would be possible to make the TEOIP become, besides a tool to organize the information that came from the police intelligence's operations, a system of alert of committal or the preparation of certain action by the criminal organization. Besides that, the study has allowed to evidence that the practices adopted by the police intelligence in the combat of criminal organizations claim for the composition of an interdisciplinary theoretic base and that the contributions consolidated in the field of Information Science were very important to the investigations, notably in the scope of organization and use of information and knowledge. From this perspective, the IS can offer theoretical and methodological contributions, keys to the comprehension of criminal action as a dynamic informational phenomenon and aggravated by the sensibility of the actions that make up its characterization, monitoring and combat.

Keywords: Police Intelligence, thesauri - methodology, sensible information.

1 INTRODUÇÃO

As operações de inteligência policial (OIP) se constituem essencialmente em complexos informacionais. Durante o desenvolvimento de uma OIP, articulam-se inúmeros desafios ligados à informação, advindos principalmente do crescimento exponencial dos dados disponíveis, que turvam o ambiente e dificultam, ou até mesmo impossibilitam, a conclusão dos trabalhos de combate a organizações criminosas.

O contexto em que ocorre a atividade de inteligência policial é extremamente dinâmico, envolto em desinformação e, ainda, cercado pelo sigilo, fatores que aumentam o grau de dificuldade para se alcançar o objetivo pretendido.

Esse ambiente dinâmico e sensível requer metodologias para tratar as informações geradas no decorrer das operações integrando documentos tradicionais, documentos eletrônicos e fluxos informacionais produzidos no curso de ações criminosas.

Essa última, o grande desafio, pois exige a implementação de instrumentos documentários capazes de orientar a indexação em tempo real de toda e qualquer ação de determinada organização criminosa (Orcrim), de maneira a que a partir dessa intermediação seja possível recuperar os assuntos de interesse. Para tanto, é necessário indexar os fluxos informacionais ocorridos no desenrolar da operação de inteligência policial.

A ciência da Informação (CI), mais especificamente nos aspectos referentes à organização e uso da informação, permeia as ações desenvolvidas no contexto acompanhado, entretanto estas ações precisam receber também uma reflexão específica da atividade de inteligência policial. Para a CI, a garantia literária, a garantia de uso e a garantia estrutural são bases para a construção de um tesouro (MOREIRA; MOURA, 2006). Todavia, no caso de sua aplicação na atividade de inteligência policial, foi preciso incorporar novos elementos que dão garantias, mas que não estão repertoriados na literatura da CI.

Assim, a meta estabelecida foi a de criar um instrumento útil para organizar e recuperar informações registradas e aquelas ainda difusas nos contextos de atuação das operações policiais.

Diante das opções de instrumentos intelectuais para organização e uso da informação, concluiu-se que o tesouro apresentava características que atenderiam às demandas colocadas, já que sua origem está relacionada com a necessidade de manipulação de grande quantidade de documentos especializados de determinadas áreas que exigem maior sistematização para a recuperação da informação (MOREIRA; MOURA, 2006).

Estruturalmente, o tesouro “é um vocabulário controlado e dinâmico de termos relacionados semântica e genericamente cobrindo um domínio específico do conhecimento” (UNESCO, 1976, p. 5). Sua principal característica é a presença de relações associativas e de equivalência entre os termos e sua finalidade é promover a coerência na indexação de conteúdos de objetos, especialmente para sistemas de recuperação e armazenamento da informação pós-coordenados, podendo ser usado também na recuperação de conteúdos de objetos em pesquisa de texto livre (ANSI/NISO Z 39.19: 2005).

Diante desse cenário, e tendo como *corpus* a Operação de Inteligência Policial Carga Suja²¹³, idealizou-se o Tesouro Eletrônico de Operações de Inteligência Policial (TEOIP) capaz de abarcar o contexto apresentado.

Durante o desenvolvimento do TEOIP, através de uma perspectiva multidisciplinar, inferiu-se que o tesouro, além de se prestar à organização das informações, também poderia servir como um sistema de alerta para a atividade policial.

Assim, alicerçada na Ciência da Informação, mas utilizando de conhecimentos advindos da Filosofia, Linguagem, Terminologia e Análise de Conteúdo, foi realizada uma análise do linguajar da organização criminosa, cujo resultado se prestou a fornecer o contexto terminológico que caracteriza e segmenta as ações criminosas. Tais termos identificados foram inseridos no tesouro.

Esses termos, quando de sua entrada no Sistema de Recuperação da Informação, teriam a capacidade de disparar um alerta sobre a preparação ou cometimento de determinada ação de interesse para a operação policial.

Com isso, além de ser uma ferramenta que apoia o trabalho policial ao tratar a informação no contexto dinâmico e sensível, o tesouro proposto ainda tem o potencial de agregar em seu interior a capacidade de alertar a equipe policial para o acompanhamento de possível ação delituosa.

O estudo e suas conclusões foram abordados detidamente em pesquisa desenvolvida junto ao PPGCI/UFMG²¹⁴. Desse trabalho, optou-se por destacar três de seus tópicos: a

²¹³ A Carga Suja prendeu, no início de setembro de 2011, 14 pessoas componentes de uma organização criminosa especializada no roubo e receptação de caminhões e cargas na região metropolitana de Belo Horizonte. Os autos das ações criminais decorrentes da operação foram colocados à disposição dos pesquisadores da Universidade Federal de Minas Gerais pela Justiça de 1ª Instância, Comarca de Pedro Leopoldo/MG, para fins de pesquisa, estudos e elaboração de dissertação de mestrado, compondo assim o *corpus* da pesquisa.

²¹⁴ Dissertação intitulada Contribuições à organização de informações dinâmicas e de aplicações sensíveis: a construção do Tesouro Eletrônico de Operações de Inteligência Policial (TEOIP),

metodologia de construção do tesouro voltado a operações de inteligência policial; a indexação proposta, tendo em vista ser uma maneira de indexação que orienta práticas, sujeita ao dinamismo das atividades inerentes às operações policiais; e o sistema de alertas como uma inovação funcional à ferramenta tesouro.

Antes, no entanto, apresentam-se sinteticamente alguns elementos que permeiam a discussão e que ajudarão o leitor a compreender melhor o ambiente em que foi pensada a pesquisa. Para isso, será abordado o contexto dinâmico e sensível das operações de inteligência policial com suas múltiplas fontes de dados, evidenciando a importância da organização e uso da informação nessas operações.

2 CONTEXTO DINÂMICO E SENSÍVEL DAS OPERAÇÕES DE INTELIGÊNCIA

As operações de inteligência policial se desenvolvem em cenários bastante dinâmicos e peculiares, envolto em uma miríade de variáveis que afetam a forma como ocorre a coleta e sistematização da informação. Nele, a atenção do profissional incumbido de tratar os dados coletados está voltada para um fluxo que corresponde a fatos sociais praticados pela organização criminosa. A indexação, em parte, é efetuada em tempo real, acompanhando o passo a passo da Ocrim.

Diante do contexto, passa-se a tratar a realidade e as etapas das operações policiais como se documentos fossem, indexando as ações. E essas ações ficariam sujeitas a reindexações em função do curso da atividade.

De modo sintético, apresenta-se a seguir algumas funções essenciais no trabalho das equipes de inteligência, que são: *coletor*, *analista* e *equipe de campo*. Esses sujeitos atuam como os “profissionais da informação” no âmbito das operações policiais.

O *coletor*, conforme o próprio nome já indica, é o policial incumbido de coletar os dados. Sua atuação é normalmente interna, nos escritórios de inteligência. Dessa maneira, vai operar, principalmente, nas estações de *Sigint*²¹⁵ (Inteligência de sinais).

Tomar-se-á como referência o exemplo de *Sigint* mais difundido pela imprensa que é a interceptação telefônica. Assim, o coletor é aquele policial responsável por ouvir e classificar os áudios decorrentes da interceptação telefônica. Todavia, na atividade policial, suas tarefas não ficam restritas à coleta, pois, devido ao fato de estar recebendo em tempo real as ligações

defendida em setembro de 2013 na Escola de Ciência da Informação da UFMG. Disponível em: <http://www.bibliotecadigital.ufmg.br/>

²¹⁵ Sigint, ou inteligência de sinais, refere-se a transmissões eletrônicas que podem ser recolhidas por navios, aviões, locais em solo, ou satélites (FBI).

telefônicas de um alvo da operação, ele ficará responsável também por acionar e orientar a equipe de campo. Dessa maneira, fica a critério desse profissional o momento do acionamento da equipe de campo quando os diálogos indicam a eminência de ocorrer alguma Situação de interesse.

No curso dessas atividades, cabe ao coletor a responsabilidade de tornar acessível, os dados importantes gerados durante o acompanhamento da referida Situação, uma vez que deverão estar disponíveis para serem recuperados por outros membros da equipe que estará incumbido de efetuar a análise e a consolidação das informações armazenadas.

Importante salientar que uma Situação é qualquer evento, criminoso ou não, que a equipe de inteligência identifique como relevante. Por isso, sua identificação, indexação e posterior recuperação se consubstanciam como uma das atividades mais importantes em uma operação de inteligência policial.

Vista a função do coletor, passamos à *equipe de campo*, também conhecida como equipe de rua, que opera em ambiente externo e desenvolve-se também como uma atividade de coleta.

Com atuação bem diversificada, uma de suas funções é o acompanhamento *in loco* das atividades da Orccrim, de modo a efetuar o registro em imagem e/ou som de determinada Situação. Ao final, sistematiza os dados coletados em um documento que descreverá o ocorrido. Normalmente é redigido um texto remetendo a fotografias, vídeos e demais recursos coletados e utilizados. Esse documento receberá um número e uma indexação²¹⁶ que o vincule aos demais documentos gerados na operação, de maneira a que, em um segundo momento, o analista possa recuperá-los em conjunto com outros que também possuam identidade com o evento.

Também é a equipe de campo que efetua levantamentos referentes à localização de pessoas físicas e jurídicas, bem como a busca por imóveis registrados em cartórios, ação importante para identificação e recuperação dos ativos frutos das práticas ilícitas perpetradas pela Orccrim. Normalmente também é o responsável pela coleta de informações com fontes humanas, conhecidas no jargão policial como informantes.

Por último, a função do *analista*. Esta se distancia um pouco da dinâmica em que estão mergulhados coletores e equipe de campo. Diferente das demais funções que trabalham

²¹⁶ A indexação é a operação que consiste em descrever e caracterizar um documento com o auxílio de representações dos conceitos contidos nesses documentos. [...] A indexação permite uma pesquisa eficaz das informações contidas no acervo documental (OLIVER, 2010, p.150).

premidas pelo tempo, uma vez que a coleta para dar um retorno adequado deve ser feita o mais próximo possível do tempo real, a análise dispõe de um período maior para sua execução.

Desenvolvendo o seu trabalho, o analista atua como um pesquisador. Ele vai utilizar os dados coletados e sistematizados pelos coletores e equipes de campo com a missão de produzir o conhecimento. Para Cepik “o ethos profissional da atividade de análise em inteligência e suas regras de produção do conhecimento são as mesmas que governam qualquer outra atividade de pesquisa” (CEPIK, 2001, p. 61).

Para aprofundar o entendimento de como se realiza o trabalho do analista, imaginemos que ele se depara com um único documento classificado pelo coletor como relevante, e que o conteúdo indexado remete a uma conversa que poderia indicar um acordo entre empresários e políticos no sentido de direcionamento de uma licitação pública.

De posse dessa informação, o analista vai efetuar pesquisas em bancos de dados que estejam à sua disposição no intuito de encontrar novos conteúdos que possam complementar aquela documentação inicial. Uma estratégia de busca possível é recuperar e escutar as conversas telefônicas que ocorreram em período aproximado da data do áudio classificado como relevante. Essa tarefa não é simples. Dependendo da operação, por vezes coletam-se centenas de áudios em um único dia, que se acumulam por semanas, meses e anos, compondo uma massa informacional de grandes proporções que se torna humanamente impossível conhecer seu conteúdo.

Nesse contexto, a indexação de conteúdo ao disponibilizar descritores específicos sobre os assuntos presentes em cada um desses áudios, possibilita uma busca por assuntos de interesse.

Assim, o analista solicitaria uma busca no sistema, por exemplo, pelo termo “licitação”. Com isso, poderia recuperar os áudios presentes naquele universo que versassem sobre o referido assunto.

De posse de um universo de áudios menor, torna-se possível escutá-los e, a partir deles, obter novos dados que somados aos anteriormente obtidos possibilitam ao analista tecer inferências a respeito da Situação. A partir daí, efetua-se nova coleta em bancos de dados disponíveis, em fontes abertas, como o Diário Oficial que porventura tenha publicado o edital da licitação, e nos demais meios que entenda pertinentes. Além disso, faz requerimentos para que a equipe de campo e os coletores empreendam novos esforços de coleta que venham a suprir os vazios informacionais em sua análise.

Feito esse apanhado geral, passa-se a um exemplo fictício que congrega o trabalho dos coletores, analistas e equipes de campo.

[parte do texto excluído por atendimento ao tamanho de artigo apresentado como Pôster]

REFERÊNCIAS

- ANSI/NISO Z 39.19: 2005 – **Guidelines for the construction, format and management of monolingual controlled vocabularies**. Bethesda, Md, NISO, 2005.
- BARDIN, L. **Análise de Conteúdo**. Lisboa: Edições 70,1977.
- BRASIL. Lei nº 12.694, de 24 de julho de 2012.
- CAMPOS, M. L. A; GOMES, H. E. Taxonomia e classificação: o princípio de categorização. **DataGramZero**, v. 9, n. 4, ago. 2008. Disponível em: <http://www.dgz.org.br/ago08/Art_01.htm>. Acesso em: 06 nov. 2011.
- CAMPOS, Maria Luiza de Almeida; GOMES, Hagar Espanha; MOTTA, Dilza Fonseca. **Elaboração de tesauro documentário** – Tutorial – Última atualização jul. 2004. Disponível em: <<http://www.conexao.org.com/bit/tesauro/index.htm>>. Acesso em 11 ago. 2013.
- CEPIK, Marco. **Serviços de Inteligência: Agilidade e Transparência como Dilemas de Institucionalização**. 2001. 310f. Tese (Doutorado em Ciência Política). Instituto Universitário de Pesquisa do Rio de Janeiro, Rio de Janeiro, 2001.
- CINTRA, A. M. M.; TÁLAMO, M.F.G.M.; LARA, M.L.G.; KOBASHI, N. Y. **Para entender as linguagens documentárias**. 2. ed. São Paulo: Polis, 2002.
- MOREIRA, Alexandra; ALVARENGA, Lídia; OLIVEIRA, Alcione de Paiva. O nível do conhecimento e os instrumentos de representação: tesouros e ontologias. **DataGramZero: revista de Ciência da Informação**, v. 5, n. 6, dez. 2004. Disponível em: <http://www.dgz.org.br/dez04/Art_01.htm>. Acesso em: 28 abr. 2013.
- MOREIRA, Manoel Palhares; MOURA, Maria Aparecida. Construindo tesouros a partir de tesouros existentes: a experiência do TCI – Tesouro em Ciência da Informação. **DataGramZero: revista de Ciência da Informação**, v. 7, n. 4, 2006.
- OLIVER, Paulo Roberto da Costa. **Projetos de ECM/BPM: Os Segredos da Construção**. São Paulo: Editora: Biblioteca 24 Horas, 2010. v. 2, 292p.
- RANGANATHAN, S. R. **Prolegomena to library classification**. 3. ed. London: 1967.
- SALDANHA, Gustavo Silva. **Micro Tesauro de Ciências Criminais**. Projeto de conclusão de curso de Biblioteconomia na escola da Ciência da Informação – UFMG. Belo Horizonte, Dez. 2005.
- UNESCO. **Unisist: Guidelines for the establishment and development of monolingual thesauri**. Paris. Dez de 1976.
- WITTGENSTEIN, Ludwig. **Investigaciones filosóficas**, Barcelona: Altaya, 1999.