



# SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO

territórios contemporâneos

ROSALI FERNANDEZ DE SOUZA

LUANA SALES

GUSTAVO SALDANHA

ORGANIZADORES



---

# **Sistemas de Organização do Conhecimento**

territórios contemporâneos

Esta publicação está disponível em acesso livre ao abrigo da licença Attribution-ShareAlike 3.0 IGO (CC-BY-SA 3.0 IGO) (<http://creativecommons.org/licenses/by-sa/3.0/igo/>). Ao utilizar o conteúdo da presente publicação, os usuários aceitam os termos de uso do Repositório UNESCO de acesso livre ([www.unesco.org/open-access/terms-use-ccbysa-port](http://www.unesco.org/open-access/terms-use-ccbysa-port)).

Esta publicação tem a cooperação da UNESCO no âmbito do projeto “Ampliação e Modernização das Ações do IBICT relacionadas às Atividades de Coleta, Armazenamento, Sistematização, Análise, Disseminação e Preservação de Dados e Informações Relativos à Ciência, Tecnologia e Inovação” (Prodoc 914BRZ2005). As indicações de nomes e a apresentação do material ao longo deste livro não implicam a manifestação de qualquer opinião por parte da UNESCO a respeito da condição jurídica de qualquer país, território, cidade, região ou de suas autoridades, tampouco da delimitação de suas fronteiras ou limites. As ideias e opiniões expressas nesta publicação são as dos autores e não refletem obrigatoriamente as da UNESCO nem comprometem a Organização.

---



#### **CONSELHO EXECUTIVO**

- › Gustavo Saldanha (Instituto Brasileiro de Informação em Ciência e Tecnologia – IBICT; Universidade Federal do Estado do Rio de Janeiro – Unirio)
- › Paulo César Castro (Escola de Comunicação – ECO/UFRJ)

#### **CONSELHO CIENTÍFICO DA COLEÇÃO**

- › Cecília Leite (Instituto Brasileiro de Informação em Ciência e Tecnologia - IBICT)
- › Miguel Ángel Rendón Rojas (Universidade Nacional Autónoma de México - UNAM)
- › Muniz Sodré (Universidade Federal do Rio de Janeiro - UFRJ)
- › Ivana Bentes (Universidade Federal do Rio de Janeiro - UFRJ)
- › Naira Christofoleti Silveira (Universidade Federal do Estado do Rio de Janeiro – Unirio)
- › Rafael Capurro (Unesco)

#### **COMITÊ CIENTÍFICO DO LIVRO**

- › Rodrigo de Salles (UFSC)
- › Tatiana de Almeida (UNIRIO)
- › Naira Christofoletti Silveira (UNIRIO)
- › Lillian Alvarez (UNB)

---

# **Sistemas de Organização do Conhecimento**

territórios contemporâneos

**Rosali Fernandez de Souza**

**Luana Sales**

**Gustavo Saldanha**

organizadores



Rio de Janeiro

2022

**Capa:** Fernanda Estevam

**Ilustração:** GK Vector (br.freepik.com)

**Projeto Gráfico:** Paulo César Castro

**Normalização e catalogação:** Selo Nyota

**Diagramação:** Fernanda Estevam

Essa obra tem o financiamento do Conselho Nacional para o Desenvolvimento Científico e Tecnológico (CNPq), da Fundação Carlos Chagas de Apoio à Pesquisa Científica do Estado do Rio de Janeiro (Faperj) e da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES).

---

S729

Sistemas de Organização do Conhecimento: territórios contemporâneos / Rosali Fernandez de Souza; Luana Farias Sales; Gustavo Silva Saldanha (org.). – Rio de Janeiro: IBICT, 2022. – (Coleção PPGCI 50 anos)

168p.

Inclui Bibliografia.

Disponível em: <https://ridi.ibict.br/>

ISBN 978-65-89167-59-4 (digital)

1. Ciência da Informação. 2. Linguagens documentárias. 3. Organização do Conhecimento. I. Souza, Rosali Fernandez de. II. Sales, Luana. III. Saldanha, Gustavo. IV. Título.

CDD 020

---



Projeto editorial em colaboração com o Programa de Educação Tutorial (PET) da Escola de Comunicação (ECO-UFRJ): Paulo César Castro (tutor) / aluno(a)s: Carolina Torres, Dandara Campello, João Maurício Maturana, Juliana Sorrenti, Kethury Santos, Lianne Henriques, Mariana da Paz, Moniqui Frazão, Robertha Braga, Sabrina Oliveira e Sara Maluf.



Programa de Pós-Graduação em Ciência da Informação (PPGCI), desenvolvido pelo Instituto Brasileiro de Informação em Ciência e Tecnologia, do Ministério da Ciência e Tecnologia e Inovação (IBICT/MCTI) em convênio com a Escola de Comunicação da Universidade Federal do Rio de Janeiro (ECO/UFRJ).

Rua Lauro Muller, 455 - 4º andar  
Botafogo - Rio de Janeiro - RJ  
<http://www.ppgci.ufrj.br>

*A pesquisa que resulta nesta publicação obteve o fomento de*

CNPq  
FAPERJ  
Capes

*& com o apoio de*

UNESCO  
IBICT  
CENACIN  
UNIRIO  
UFRJ





*À comunidade científica da área  
de Organização do Conhecimento*



*“Se me derem a sabedoria com a condição de que eu a  
guarde para mim sem poder transmiti-la, eu a  
recusarei. Não é agradável possuir um bem quando  
não podemos dividi-lo.”*

Sêneca



---

# Sumário

- 13** Apresentação  
**Rosali Fernandez de Souza, Gustavo Saldanha e Luana Sales**
- 15** Prefácio  
**Luana Farias Sales, Gustavo Saldanha e Rosali Fernandez de Souza**
- 19** Linguagens documentárias para democracia documentária  
**Gustavo Saldanha, Rosali Fernandez de Souza e Luana Sales**
- 39** A organização do conhecimento na Fiocruz: uma investigação dos seus sistemas de organização do conhecimento  
**Denizard Costa da Silva, Elaine Lucia da Silva, Karina Veras Praxedes dos Santos e Thaís de Oliveira Sant’Anna Campanhã**
- 77** Os sistemas de organização do conhecimento das universidades: um estudo de caso das bibliotecas da Universidade Federal do Rio de Janeiro  
**Carla Beatriz Marques Felipe, Eliane Azevedo Gomes, Milene Couras da Cruz**
- 111** Os Sistemas de Organização do Conhecimento do Repositório de Acesso Livre à Informação Científica (Alice) da Empresa Brasileira de Pesquisa Agropecuária (Embrapa)  
**Alessa Fabíola dos Santos Ceslinski, Fabiana Magno de Lacerda, Lorrane de Souza Saluzi, Maria Cecília Jardim Barros, Skrol Salustiano**

- 135** Sistemas de organização do conhecimento: uma análise do Tesouro Unesp  
**Carlos Victor de Oliveira, Jamille Abreu Passalini de Sousa e Kelly Maria Ayala de Carvalho**
- 149** A importância das linguagens documentárias nos ministérios para o acesso à informação no Brasil  
**Marcelle Costal de Castro dos Santos, Patrícia Conceição Romeu da Fonseca, Priscila Ramos Carvalho**

**A** ORGANIZAÇÃO DO CONHECIMENTO, PARTICULARMENTE NA CIÊNCIA DA INFORMAÇÃO e na Biblioteconomia, aborda questões históricas e epistemológicas da produção do conhecimento, assim como os processos de representação de documentos em sistemas de recuperação da informação. A singularidade de cada sistema de informação é determinada pela missão, metas e objetivos institucionais onde atuam, definindo os tipos de documentos que compõem o seu acervo visando atender as necessidades de informação dos seus usuários. Para que tal objetivo seja alcançado, é relevante e mesmo imprescindível o papel do profissional da informação nos processos de classificação e indexação. Nesse contexto, os diferentes tipos de sistemas de organização do conhecimento desempenham papel de destaque.

A complexidade deste panorama motiva e instiga professores e pesquisadores a investigar sobre os aspectos históricos e as bases epistemológicas que caracterizam os diferentes tipos de sistemas de organização do conhecimento construídos ao longo do tempo, como também no delineamento de análises de desempenho na função de organizar e representar o conhecimento.

Nesse sentido, é com imenso prazer que como docentes, Gustavo Saldanha, Luana Sales e eu, apresentamos o produto da experiência de ensino da disciplina Sistemas de Organização do Conhecimento, em capítulos, os resultados das pesquisas realizadas pelos alunos que cursaram esta disciplina por nós ministrada em 2020. As pesquisas abordam SOC em uso em ambientes acadêmicos, como institutos de pesquisa (FIOCRUZ e EMBRAPA) e universidades (UFRJ e UNESP) e, em órgãos da administração pública federal (MINISTERIOS). As diferentes abordagens metodológicas de coleta e análise dos dados adotadas pelos grupos de trabalho, revelaram resultados e discussões que acreditamos relevantes sobre o desempenho de SOC nos ambientes estudados, assim como para motivar pesquisas em outros ambientes de informação.





**P**RIMEIRO SEMESTRE DE 2020. VIVÍAMOS UM MUNDO NOVO. TODOS ISOLADOS por causa da pandemia de COVID-19 que ameaçava dizimar parte da população mundial. Escolas, Universidades e Programas de Pós-Graduação precisavam se reinventar, adaptando-se ao novo contexto, descobrindo e redescobrimo ferramentas e novas metodologias de ensino que permitissem aos estudantes absorver o máximo de conhecimento possível diante de tal cenário de isolamento.

Foi nesse contexto que ministrávamos a disciplina “Sistemas de Organização do Conhecimento” no ano do cinquentenário do PPGCI IBICT UFRJ. Ao mesmo tempo em que tentávamos nos reinventar, diante de um mundo de incertezas, discentes buscavam se dedicar, concentrando-se na assimilação de todo o conteúdo ministrado remotamente, ao mesmo tempo em que se dividiam nesse novo ambiente, em uma série de papéis que até então eram exercidos separadamente. Esses discentes eram paralelamente, estudantes, profissionais, mães, pais, filhos, esposas e esposos.

O objetivo da disciplina era o estudo de “princípios teóricos e metodológicos de elaboração de linguagens de representação e recuperação da informação”. Tipologias de sistemas de organização do conhecimento e representação de informação. Análise de desempenho de esquemas e tabelas de classificação, taxonomias, tesouros, vocabulários controlados, ontologias, terminologias e glossários em função de usos específicos. A moderna concepção da participação das comunidades virtuais colaborativas”, conforme a ementa divulgada.

Considerando as limitações de tempo e espaço, o conhecimento e a possibilidade acesso tecnológico de alguns, esses discentes foram mais que vitoriosos, fazendo-se presente nas aulas, revendo a gravação quando, por algum motivo precisavam se ausentar, executando as tarefas propostas, lidando com o vírus que, para uns parecia fictício, mas que para outros era mais do que real, levando parentes e amigos.

A disciplina se materializou à medida em que além de ensinar, os desafiávamos a se aventurarem na Web, buscando sistemas de Organização do Conhecimento

que pusessem ser analisados empiricamente, relacionando o que era discutido teoricamente através das aulas remotas

Ao final da disciplina, um novo desafio: apresentar esses trabalhos em seminários de aula e transformá-los em um potencial capítulo para o livro que agora vocês leem aqui.

A organização desse livro é uma conquista que precisa ser registrada como ato de resistência daquelas vidas que, mesmo desacreditadas, com medo, isoladas, lutaram bravamente para não serem engolidas pelo sistema, virando mais um número diante do macabro cenário de morte que vivemos nos últimos anos. É um ato de resistência também, porque diante de tudo isso, foram pessoas que continuaram suas pesquisas perante toda a crítica à ciência no negacionismo em crescimento. E são essas guerreiras e guerreiros que obstinadamente se tornaram autorias desse livro que organizamos.

O primeiro capítulo revela a motivação e o direcionamento que a organização da disciplina se deu de forma que pudéssemos também obter contribuições para o projeto de pesquisa que iniciávamos conjuntamente naquele momento. Este projeto intitulado Hagar: Linguagensdocumentárias.br tem como objetivo a identificação, avaliação, preservação, integração, criação, multiplicação e disponibilização das linguagens documentárias (catálogos, classificações, glossários técnico-científicos, classificações, vocabulários controlados, tesouros, taxonomias, ontologias, dentre outras) produzidas no Brasil com foco na inovação social e econômica para o desenvolvimento do país. O objetivo do Projeto Hagar é a criação de um repositório nacional de sistemas de organização do conhecimento (SOC) que possam ser consultados, acessados e ter seus conceitos e sua estrutura reusados em diversos contextos informacionais.

Neste sentido, à medida em que os discentes levantavam os SOC, nós também reuníamos material para ser agregado ao repositório, que denominamos Repositório HAGAR – uma pequena homenagem, aquela que iniciou muitos de nós docentes do PPGCI e do estado do Rio de Janeiro na temática “Organização do Conhecimento”

Do capítulo dois em diante, temos os relatos das pesquisas de nossos discentes que encontraram suas próprias metodologias de análise e desenvolvimento do texto.

No capítulo 2, encontramos um levantamento e análise dos sistemas de organização do conhecimento das unidades da Rede de Bibliotecas da FIOCRUZ localizadas no Rio de Janeiro, além de uma biblioteca virtual e o repositório institucional da Fundação: o ARCA. Esta pesquisa resultou na análise de 13 SOC que poderão ser registrados no Repositório HAGAR, caso atendam os critérios estabelecidos a priori.

O capítulo 3 faz uma análise dos Sistemas de Organização do Conhecimento utilizados no repositório de acesso livre da EMBRAPA, denominado ALICE. A análise

desse repositório identificou o uso de dois SOC, além dos termos livres sugeridos como palavras-chave pelos autores das publicações contidas no ALICE, sendo eles: o Thesagro - Thesaurus Agrícola Nacional e o *National Agricultural Library's Thesaurus and Glossary* (NAL Thesaurus) da Biblioteca Nacional de Agricultura dos Estados Unidos. Apesar da identificação dos dois SOC, apenas o Thesagro foi analisado, dado o objetivo do trabalho proposto que era levantar e analisar SOC para compor o repositório nacional HAGAR.

No capítulo 4, os autores escolhem o contexto da Universidade Federal do Rio de Janeiro (UFRJ) para fazer o levantamento dos Sistemas de Organização do Conhecimento utilizados nas diversas bibliotecas que compõem o Sistema de Bibliotecas (SIBI). Por meio de entrevista com bibliotecários responsáveis por um quantitativo considerável de bibliotecas e centros de informação da UFRJ, foram identificados 15 SOC em uso no SIBI. Os autores do capítulo 5, por sua vez, embora também tenham escolhido trabalhar com SOC de bibliotecas universitárias, escolheram uma universidade de outro estado: a UNESP que, diferentemente do exemplo anterior, para promover a comunicação entre as bibliotecas de seus diversos campi espalhados pelo estado de São Paulo, se vale de um SOC único que padroniza a linguagem de indexação utilizada em todas as bibliotecas. Neste caso, puderam se debruçar com mais profundidade na análise feita com base nos critérios de avaliação de tesouros disponibilizados no site da Professora Hagar Espanha Gomes: o BITI: Biblioteconomia, Informação e Tecnologia da Informação.

O capítulo 6 apresenta um levantamento das linguagens documentárias utilizadas em âmbito governamental, com um recorte feito pelos Ministérios do atual governo. Através desse estudo, os autores identificaram 33 sistemas de organização do conhecimento em uso nos diversos Ministérios. O trabalho, porém, não investigou os vocabulários utilizados pelas instituições governamentais ligadas aos ministérios, o que poderia ter levado ao reconhecimento de um número ainda maior de SOC usados no governo.

Com isso, convidamos vocês, leitores, a conhecer um pouco do percurso da disciplina que se inicia no capítulo a seguir com uma apresentação detalhada dos objetivos, métodos e teorias que embasam o Projeto Hagar e se encaminha logo após para os relatos de nossos discentes em um misto de teoria e empiria.



# Linguagens documentárias para democracia documentária

Gustavo Saldanha<sup>1</sup>, Rosali Fernandez de Souza<sup>2</sup> e Luana Sales<sup>3</sup>

---

## 1 Introdução

APRESENTA-SE AQUI O PERCURSO INICIAL DO PROJETO “HAGAR: LINGUAGENS DOCUMENTÁRIAS PARA DEMOCRACIA DOCUMENTÁRIA”, caracterizado como pesquisa científica, teórica e empírica. A proposta, preliminarmente representada pelo desenho-piloto do projeto “Linguagens documentárias fluminenses para a inovação: organização do conhecimento para ciência, cultura e sociedade no Estado do Rio de Janeiro”, submetido ao edital de pesquisa científica apresentado à Chamada Jovem Cientista do Estado – Faperj por Gustavo Saldanha (209-2021), coloca em questão as relações socio-cognitivas da organização do conhecimento, com foco nos potenciais das ações do conhecer visando a inovação pela via das linguagens documentárias. Tais relações remontam a fundamentação do campo de estudos da informação, seja este pensado como construção histórica de longa extensão, diacrônica, aquém do século XX, seja este tomado como resultado de uma análise sincrônica, formalizado no contexto contemporâneo de desenvolvimento das tecnologias da informação e da comunicação (TICs), em ambos os casos sob a missão de tratar dos/os documentos e metadocumentos produzidos pelas sociedades.

As linguagens documentárias são aqui reconhecidas a partir de instrumentos como classificações, terminologias, vocabulários controlados, tesouros, ontologias. São todos estes elementos mediadores entre sujeitos, artefatos e processos que permitem, pela via da linguagem, a identificação, a recuperação e o uso da informação, constituindo, ao mesmo tempo, contextos para descoberta de novos conhecimentos a partir do entrecruzamento de linguagens. Esta pesquisa parte do pressuposto da condição fundacional das linguagens documentárias de permitir o acesso aos

---

1 Doutor em Ciência da Informação – PPGCI IBICT UFRJ; pesquisador titular (IBICT); professor adjunto (UNIRIO); bolsista de produtividade CNPq 2; bolsista jovem cientista do estado Faperj.

2 Doutor em Ciência da Informação – PPGCI IBICT UFRJ; pesquisadora titular (IBICT)

3 Doutor em Ciência da Informação – PPGCI IBICT UFRJ; pesquisador titular (IBICT); professor adjunto (UNIRIO); bolsista de produtividade CNPq 2; bolsista jovem cientista do estado Faperj.

saberes, bem como a capacidade de potencializar o desenvolvimento de conhecimentos inovadores.

A inovação é caracterizada, aqui, a partir do Manual de Oslo da Organização para Cooperação Econômica e Desenvolvimento (OCDE, 1997), ou seja, como a efetiva introdução de um dado bem ou um dado serviço novo, com significativa condição de aperfeiçoamento em relação às características ou usos previstos, incluindo também a implementação de novos métodos ou processos de produção, distribuição, marketing ou organizacionais novos ou significativamente melhorados.

O pressuposto estrutural indicado está, pois, em articular as políticas da OCDE a partir das linguagens documentárias, acreditando-se que tais instrumentos são fundamentais para o foco no desenvolvimento econômico sustentável e de emprego, a ampliação da qualidade dos padrões de vida e a estabilidade financeira, dentro do contexto de expansão de uma economia estável, dado que todos os setores da sociedade dependem de classificações operando com qualidade e dinâmica para a realização de seus resultados.

Embora as classificações sejam fundamentais para a operação de resultados, as classificações, bem como as demais linguagens documentárias existentes encontram-se dispersas, dificultando sua encontrabilidade, acesso, uso e reuso na estruturação de sistemas e conseqüentemente a encontrabilidade, o acesso, uso e reuso também das informações armazenadas nesses sistemas.

Conjectura-se que a identificação e a integração de linguagens documentárias em sistema único, aberto, público, colaborativo e orientado para democracia documentária, pode aumentar significativamente as possibilidades de inovação no campo informacional e da organização do conhecimento, impactando os mais diversos setores da sociedade, como já mencionado.

No âmbito deste projeto trata-se de compreender o desenvolvimento de uma teoria da informação, com foco na organização do conhecimento, orientada para a visibilidade e a transformação das condições socioeconômicas a partir dos sistemas informacionais. A construção política das ações do classificar é, deste modo, compreendida aqui como força para a fundamentação de condições de desenvolvimento de um dado contexto. Para tal, o projeto reúne parcerias em um plano interdisciplinar e interinstitucional orientadas para uma apreensão da representação dos saberes elaborados por diferentes esferas institucionais, tendo o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), casa bibliográfica por missão, desde sua criação em 1954, como eixo de articulação das redes de metalinguagens.

A questão de pesquisa é enunciada pela indagação: como podemos ampliar os potenciais de identificação e de integração das linguagens que instituem a inovação

através dos sistemas informacionais? A hipótese geral que atravessa toda a reflexão aponta que, para atingir as margens de solução da questão estamos diante de um dilema da linguagem, espaço-tempo onde é depositado todo conjunto diverso e dinâmico dos saberes; logo, um método que pressuponha o papel filosófico, social e instrumental da linguagem sustentará a argumentação em busca dos resultados para a indagação da inovação. As hipóteses secundárias apontam para a) o construto simbólico como estrutura para uma fundamentação teórica e aplicada para tal representação de sujeitos, artefatos e processos e b) o papel inovador da organização do conhecimento como centralidade no desenvolvimento de estruturas transversais voltadas para as categorias do conhecimento. A procura da pesquisa está, deste modo, em identificar, reunir e analisar a produção de linguagens documentárias em território nacional com foco na inovação, a saber, a construção de conhecimento com impactos compartilhados entre ciência, cultura e sociedade.

O projeto tem a sua sede no Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), Rio de Janeiro, em articulação e cooperação com a Universidade Federal do Estado do Rio de Janeiro (UNIRIO) e a Universidade Federal do Rio de Janeiro (UFRJ). A partir do elo destas três instituições, parte-se inicialmente da parceria teórico-empírica com Instituto Brasileiro de Geografia e Estatística (IBGE), Fundação Biblioteca Nacional (FBN) e Fundação Casa de Rui Barbosa (FCRB), tendo como horizonte as redes de instituições informacionais.

## **2 Justificativa**

A aliança entre organização do conhecimento e inovação colocada aqui sustenta-se como um elo fundamental para o desenvolvimento socioeconômico de cada contexto, sendo o cenário brasileiro o nosso foco. Deste modo, a proposta de pesquisa se justifica a partir das seguintes dimensões: econômica, política, social e científica.

As dimensões da justificativa ficam assim delineadas:

- Econômica: a perspectiva da capacidade de recuperação da informação e da descoberta de novos conhecimentos potencializada pelas linguagens documentárias confere ao projeto as margens de contribuição para o desenvolvimento econômico do país;
- Política: na medida em que permite revelar os modos de nomeação e de uso de artefatos e de processos, de sujeitos e de instituições, as linguagens documentárias são fundamentais para a tomada de decisão governamental, sendo, pois, o projeto parte de um potencial de identificação e de construção, a partir da ordenação, de fontes para políticas públicas;

- Social: o poder de visibilizar sujeitos, práticas e contextos propiciado pelas linguagens documentárias permite a construção das margens de compreensão da dinâmica de um povo e de suas relações, sendo, pois, o projeto, um canal fomentador de fontes para o espelhamento dos saberes no Brasil;
- Científica: a dinâmica de ordenação e de visibilidade das linguagens documentárias produzidas no Brasil permitirão a compreensão do vasto conjunto de vocabulários e de correlações que darão sustentação ao desenvolvimento de políticas científicas e atenderão, ao mesmo tempo, diretamente aos pesquisadores interessados nos mais diferentes domínios da pesquisa no país.

Do ponto de vista das condições de realização do projeto, justifica-se ainda o potencial de realização a partir das condições de produção bibliográfica, redes de pesquisa, relevância da área do conhecimento e experiências progressas de investigação teórico-empírica e de captação advindas da construção histórica da organização do conhecimento desenvolvidas no IBICT.

O projeto parte centralmente no âmbito dos estudos epistemológicos, históricos e teóricos da Ciência da Informação, bem como nas experiências teóricas e empíricas da Organização do Conhecimento, sob via das teorias críticas orientadas para a organização dos saberes, ou seja, para o território da pluralidade, da diversidade e da luta social das distintas formas do conhecer. Integram, pois, a estrutura do o projeto, os grupos de pesquisa

- “Ecce Liber: filosofia, linguagem e organização dos saberes”, envolvido com um conjunto de publicações e de ações que colocam a transversalidade das dinâmicas temáticas acima indicadas, sob a liderança de Gustavo Saldanha;
- “Organização do conhecimento”, que explora a disciplinaridade e interdisciplinaridade de classificações na estruturação e nomeação de diferentes áreas do conhecimento, sob a liderança de Rosali Fernandez de Souza;
- “BRIET: Biblioteconomia, Representação, Interoperabilidade, E-Science e Tecnologias”, sob a liderança de Luana Sales e Luís Sayão– que vem desenvolvendo diversos estudos no âmbito da gestão e curadoria de dados e informações digitais, em um diálogo constante com Organização do Conhecimento, especialmente em seus aspectos representacionais.

O percurso que leva à construção deste projeto, à configuração de suas parcerias e ao seu enfoque nos laços entre organização do conhecimento e inovação contém um conjunto de experiências que permitiram a definição de seu atual arranjo.



Tais experiências podem ser divididas em teóricas e empíricas. As experiências teóricas reúnem os construtos acadêmico-científicos, perpassando:

Experiências teórico-acadêmicas: informação, mediação e organização do conhecimento em perspectiva

- Colóquios de Organização do Conhecimento (COC – IBICT): o COC representa um diálogo regional constituído em 2015 por Rosali Fernandez de Souza e Gustavo Saldanha, e desenvolvido junto de Luana Sales, com foco nas múltiplas possibilidades de reflexão e de aplicação das teorias e dos métodos de organização do conhecimento, principalmente no plano sociopolítico do domínio. A motivação dos COC é apresentar e discutir projetos de pesquisadores e estudantes de programas de pós graduação em Ciência da informação contribuindo para sua divulgação e registro. As quatro primeiras edições dos COC se realizaram em 2015, 2016, 2017 e 2018. Somando esforços em favor da Organização do Conhecimento, na edição de 2019, o V COC foi realizado em conjunto com o Fórum de Organização do Conhecimento (III FOCO) da UNIRIO no conagraçamento de pesquisas propiciando a reflexão sobre temas que perpassam a relação da Biblioteconomia e da Ciência da Informação com a Organização do Conhecimento.

A constatação da diversidade de temas da organização do conhecimento em suas abordagens epistemológicas e práticas, a ampliação de parceiras com a inclusão da UFF, e a constatação de um número cada vez maior no número de participantes, foram as principais motivações para uma ainda maior ampliação do evento, o que se concretizou na realização de um Seminário de Organização do Conhecimento (SEMOC) em maio de 2021. A pandemia impediu a realização da edição do evento em 2020. No entanto, a comunidade de organização do conhecimento resistiu. Durante o ano de 2020, ano coincidente com o cinquentenário do PPGCI IBICT, o SEMOC foi planejado tendo como abordagem central a história da organização do conhecimento no país. Em quatro dias de atividades, contando com a participação de conferencistas, nacionais e do exterior, de representantes das instituições parceiras na coordenação de mesas redondas e de palestrantes de instituições acadêmicas das diferentes regiões do país, trajetória histórica da organização do conhecimento no país nos últimos 50 foi contada. Os relatos apresentados, evidenciaram a representatividade e o amadurecimento da Organização no Conhecimento na Ciência da Informação e na Biblioteconomia no país. O panorama revelado amplia como ainda mais relevantes e promissoras as perspectivas de pesquisa e de ensino da Organização do Conhecimento para pesquisadores e estudantes.

- As Disciplinas ‘Organização do Conhecimento’ e ‘Sistemas de Organização do conhecimento’ no PPGCI IBICT, com diferentes nomeações, sempre estiveram presentes na trajetória dos 50 anos da pós graduação. Ao longo do tempo, ementas, conteúdos programáticos, definição dos tópicos considerados fundamentais e a atividade de avaliação dos alunos foram se modificando, modernizando e adaptando, buscando acompanhar o fluxo natural do avanço da organização do conhecimento em seus aspectos epistemológicos e práticos no âmbito da Ciência da Informação e da Biblioteconomia no país e no mundo. A disciplina ‘Organização do Conhecimento’ em linhas gerais discute, em bases filosóficas e epistemológicas o conceito de conhecimento e o conceito organização do conhecimento, nos diferentes contextos de produção e uso de informação. A reflexão está no que significa organizar o conhecimento e para que/para quem. A disciplina ‘Sistemas de Organização do Conhecimento’ busca caracterizar os diferentes tipos de SOC, discutindo funções e metodologias de construção e de análise de desempenho, refletindo sobre as características de cada SOC. Em cada uma das disciplinas tratamos das questões epistemológicas e práticas da organização do conhecimento, não como mutuamente exclusivas, mas sim mostrando a natureza interativa entre ambas.

Dois exemplos objetivos dessa configuração acadêmico-científica estão a) na publicação, em 2017, do Dossiê “Organização do Conhecimento & Gênero”, pela Revista Informação & Informação, da Universidade Federal de Londrina, com foco nas teorias críticas de gênero e o dilema da representação da mulher nos sistemas de organização do conhecimento; b) no dossiê do periódico LIINC EM REVISTA, do IBICT em parceria com a UFRJ, “Organização Conhecimento: dilemas sociopolíticos e suas lutas históricas”, com chamadas no primeiro semestre do ano corrente e previsão de publicação para novembro de 2018, publicação esta orientada para a reunião das mais diferentes tendências de pesquisa em teoria da classificação;

- Fórum Internacional A Arte da Bibliografia: o fórum colocou em suspeição, a partir de 2014, o potencial crítico-conceitual da noção de “bibliografia”, tanto em seu percurso histórico como em suas aplicações contemporâneas, demonstrando a vitalidade teórico-metodológica desta categoria. A partir desta exploração, foi possível remontar, até o momento, um vasto número de potencialidades de mapeamento do conhecimento e de sua análise pela via bibliográfica. Até o momento, o fórum já realizou 4 (quatro) eventos, os Seminários Internacionais A Arte da Bibliografia, resultando em 4 (quatro)

dossiês publicados nos anos de 2015, 2016, 2017 e 2018, respectivamente, nos periódicos científicos da Área Comunicação e Informação, Informação & Informação, InCID, Perspectivas em Ciência da Informação, Informação & Informação

- Rede Franco-Brasileira de Pesquisadores em Mediação e Usos Sociais dos Saberes e da Informação (Rede Mussi): completando em 2018 seus 10 (dez) anos de experiência teórico-metodológica, sob a liderança de Regina Marteleto pelo Brasil e Viviane Couzinet pela França, o coletivo de pesquisadores de Brasil, França e de outros países, vinculados aos grupos que compõem o quadro, a Rede Mussi produziu 7 (sete) eventos internacionais em sua trajetória, explorando o conceito de mediação e as múltiplas possibilidades de aplicação do mesmo nos territórios sociais; interessa ao projeto o arcabouço teórico que permite o diálogo entre mediação, representação social e experiências em saúde, como encontrado nos trabalhos de Regina Marteleto.
- Participação efetiva em eventos nacionais e internacionais da área de Organização do Conhecimento, com apresentação de trabalhos e publicações, como por exemplo os eventos da ISKO internacional, ISKO – Brasil, ISKO ibérico, nos anos de 2018, 2019 e 2021.
- Participação em movimentos nacionais e internacionais em prol da Ciência Aberta e da boa gestão de dados de pesquisa, como o 4º Plano de ação nacional brasileiro da Parceria para o Governo Aberto, liderado pela Open Government Partnership (OGP), em que, durante os anos de 2018-2020 participamos no Compromisso 3, cujo objetivo era estabelecer mecanismos de governança de dados científicos para o avanço da Ciência Aberta no Brasil; o Research Data Alliance - RDA e GO FAIR – esses dois últimos com a participação na criação de um nó brasileiro, tornando o Brasil mais representativo nos cenários internacionais, especialmente no que diz respeito às suas pesquisas em Ciência da Informação e Organização do Conhecimento.

Experiências empíricas: dos sistemas informacionais aos sistemas para inovação

O Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) e seus sistemas informacionais para a ciência e a sociedade: o IBICT conta com uma experiência teórica e empírica de 64 (sessenta e quatro) anos com foco na responsabilidade pública de desenvolvimento de sistemas e de projetos no domínio informacional e metainformacional. Dentre os exemplos, podemos destacar a Biblioteca Digital de Teses e Dissertações (BDTD) do Brasil e o Canal Ciência, maior portal de divulgação científica do país. Além disso, destaca-se a customização do software D-Space para o uso e reapropriação de revistas científicas no país e no contexto

internacional, bem como o sistema de informação sobre o cerrado brasileiro. Do ponto de social, cultural e ambiental, demarca-se as ações do IBICT como:

- Saberes do Cerrado: projeto, coordenado pelo IBICT em parceria com Jardim Botânico de Brasília e Universidade Federal de São Carlos, que visa à valorização social da biodiversidade do bioma cerrado e ao reconhecimento atual do papel atribuído às áreas protegidas, considerando a diversidade de saberes associados à transformação ou à preservação das paisagens do Distrito Federal;
- Mapa da inclusão digital: projeto que oferece um cenário, permanentemente atualizado, a respeito dos resultados das iniciativas de inclusão digital no país, considerando três níveis de atuação: Governo (federal, estadual e municipal), Setor Privado e Terceiro Setor.
- Revista Ciência da Informação: primeira revista de América Latina e Caribe dedicada à pesquisa em Ciência da Informação;
- Revista científica Inclusão Social: periódico eletrônico semestral que publica trabalhos no âmbito da inclusão social, com temas ligados a ações, programas, projetos, estudos e pesquisas voltados à problemática da inclusão dos cidadãos na sociedade da informação.

### **3 Dos objetivos**

Como descrito, esta pesquisa, contextualizada nas discussões contemporâneas organização do conhecimento com foco na inovação, realiza um estudo teórico e empírico no contexto da organização das linguagens produzidas por e para o Brasil, a partir do diálogo entre linguagens, tecnologias, instituições e integração pragmática de dados. Diante das questões previamente apresentadas, a proposta de investigação em pauta tem como objetivo geral “Investigar, identificar e mapear as experiências constituídas e em curso de organização do conhecimento no Brasil a partir de redes sociotécnicas de coleta, análise e organização de dados de linguagens documentárias desenvolvidas por instituições públicas e privadas no Brasil” e como objetivos específicos: a) Identificar e analisar teorias e metodologias em organização do conhecimento com foco em dinâmicas científicas, sociais e culturais; b) Mapear, reunir e ordenar a produção do conhecimento sobre linguagens documentárias no Brasil; c) Aplicar métodos e técnicas da organização do conhecimento no corpus documentário; d) Desenvolver instrumentos para uso e reuso de metadados na interface de linguagens documentárias.

### **4 Métodos**

A construção do marco teórico da proposta atravessa o contexto de discussão da Ciência da Informação em direção à Organização do Conhecimento, chegando

até o debate sobre as configurações teórico-metodológicas para o olhar sobre o olhar da inovação. Nesse percurso, discutimos a “construção pragmática das ações de linguagem” e a análise de domínio, conduzindo-nos a uma organização do conhecimento com foco na inovação.

A operacionalização geral das etapas de coleta da pesquisa se dá a partir de três perspectivas metodológicas: bibliográfica-documental, teórico-institucional e desenvolvimento de redes. O passo inicial do projeto se dá a partir da identificação, do cotejamento e do estudo de fontes bibliográficas referentes ao domínio, partindo de experiências já constituídas. A segunda etapa procura correlacionar o papel da linguagem e das instituições informacionais (bibliotecas, arquivos e museus), com vistas a identificação de centros de coleta, processamento e memória de dados. A terceira etapa visa a integração pragmática de dados em um portal na web dedicado à política de visibilidade da produção de saberes com vistas à inovação no Brasil.

- Base teórico-conceitual do método: a análise de domínio no caminho da inovação

A análise de domínio se constitui atualmente como um dos principais construtos teórico-metodológicos da Ciência da Informação. Trata-se de parte de um conjunto de abordagens que despontaram nos anos 1990 no campo a partir de uma teoria do conhecimento via linhagens pragmáticas. Dentre os representantes desse movimento, podemos identificar Frohmann (1992), Brier (1996), Marteleto (1994, 1995, 1996) e o pensamento enunciado por Hjørland & Albrechtsen (1995).

Hjørland & Albrechtsen (1995) lançaram, em *Toward a new horizon in information science: domain-analysis*, apresentando a Análise de Domínio como um programa teórico-metodológico para a Ciência da Informação. O foco está na compreensão de determinados domínios (coletivos envolvendo sujeitos, instituições, ações, linguagem) que são constituídos por comunidades discursivas.

O tripé teórico da Análise de Domínio seria a abordagem fundamentalmente social, abrangendo uma psicologia social, uma sociolinguística, uma sociologia do conhecimento e uma sociologia da ciência; uma abordagem funcionalista, no sentido de que se propõe operacionalizar construtos teóricos em formas explícitas de informação e de comunicação; uma abordagem filosófica de natureza realista, com foco no real interpretado como forma de construção subjetiva dos sujeitos no mundo social. (HJORLAND; ALBRECHTSEN, 1995)

A proposta de Hjørland & Albrechtsen (1995) não é a-histórica. Ela dialoga, em uma epistemologia diacrônica, com diferentes teorias constituídas no campo infor-

macional, demarcando o papel da bibliografia e da classificação bibliográfica, por exemplo, em percursos teóricos como aqueles de Henry E. Bliss e S. R. Ranganathan. No desenvolvimento da teorização da Análise do Domínio, alguns elementos se estabelecem como fundamentais para a compreensão do foco deste projeto no âmbito da organização do conhecimento, a saber, a relação entre organização do conhecimento e inovação.

Primeiramente, Hjørland & Albrechtsen (1995) definem como marco inicial da Análise de Domínio o fato de que o homem é primeiramente um sujeito, vivendo e agindo em um mundo bio-físico, sócio-cultural e subjetivo. Em segundo lugar, viver e agir constitui o a priori do conhecimento, ou seja, é a ação informacional dos sujeitos que constitui seus modos de vivência. Em terceiro lugar, desde que viver e agir constitui o a priori do conhecimento, o conhecimento é constituído de modo que um meio para a aplicação do conhecimento bem construído diretamente ou indiretamente servirá para viver e agir. Assim, os saberes constituídos por cada comunidade discursiva são, justamente, os modos de vida, as formas de ação do próprio sujeito inserido neste coletivo. Esta questão é demarcada na quarta condicionante da Análise de Domínio, ou seja, quando o conhecimento se torna parte de um sistema de ação, ele funciona como um determinante interno de ação.

Como quinto princípio, Hjørland & Albrechtsen (1995) determinam que há uma interação contínua entre conhecimento e ação assim que o conhecimento é criado até o fim da ação, assim como experiências que o ator adquire diretamente da ação influenciam a ação subsequente. Assim, o conhecimento provém de uma informação que é construída ao longo da ação dos indivíduos em seu cotidiano, em sua práxis. Assim, as interações, as trocas sociais, a formulação de linguagens para dinamizar a comunicação, os jogos de linguagem, tudo isso é objeto de investigação da CI – e está muito aquém da informação. A sexta condicionante estabelece que o conhecimento de valor, o conhecimento factual e o conhecimento procedimental são três tipos de conhecimento conectados aos três tipos de determinantes internos da ação. O conhecimento de valor significa saber o que cumpre os critérios dos bons valores. O conhecimento factual significa ter convicções (crenças) verdadeiras (legítimas) sobre os três mundos em que uma ação é viva. O conhecimento procedimental significa saber como alcançar um ato específico ou ato seqüencial. O sétimo princípio determina que o conhecimento pode ser inarticulado ou articulado. Conhecimento inarticulado é, por instância, conhecimento tácito, familiarizado, conhecimento do conhecimento. O conhecimento pode ser articulado em linguagem, ciência e arte. Assim, o conhecimento entre aquele que apenas se realiza na atmosfera dispersa da comunicação, não formalizado, isto é, uma espécie de meta-conhecimento, ligada a informações que se sedimentam em tradições no

tempo; e aquele conhecimento formalizado, que é costurado através de uma sintaxe e legitimado por um vocabulário, por um léxico que sustenta a comunicação.

O resultado do percurso de apresentação a Análise do Domínio nos permite tomar a linguagem como um objeto preliminar de estudo. Por isto, todo o conhecimento é, antes, um complexo de discursos compartilhados. Em outras palavras, todas as formas de conhecimento são narrativas sob a sombra de uma cultura específica. E o modo de conhecê-la (cada uma destas culturas) está em uma compreensão pela via pragmática da linguagem.

- Base teórico-filosófica do método: a pragmática das ações da linguagem

A abordagem pragmática, para além da linhagem sintática e da linhagem semântica, está interessada nos modos como os significados são culturalmente produzidos, depositados e rearticulados, dinamicamente, na linguagem. A influência desta corrente no campo informacional ganhará diferentes apropriações, como já indicado, percorrendo abordagens pós-estruturalistas até abordagens dialéticas de compreensão do papel da organização do conhecimento. Tais abordagens foram avaliadas por Rendón Rojas (1996), González de Gómez (1996), compreendendo as mutações no plano epistemológico e teórico-aplicado da Ciência da Informação naquela década. No mesmo período, Jacob & Albrechtsen (1999) apresentaram uma síntese da configuração conceitual e empírica da filosofia da linguagem de ordem pragmática nos estudos informacionais.

Dentre as abordagens pela via da teoria pragmática identificadas por Jacob & Albrechtsen (1999) estão o antidogmatismo e o *fallibilism* que aborda o argumento que vai contra a rigidez do racionalismo científico. Os princípios pragmáticos permitiriam entender o mundo em sua instabilidade, o mundo como entidade fadada à descontinuidade, ou seja, à contingencialidade, considerando os indivíduos como atores que convivem com o imprevisto e que conjugam com seu grupo social o conhecimento que vivenciam e reconstróem, evidenciando assim o pluralismo da linguagem cotidiana, outra lição da pragmática.

Em termos pontuais, o antifundamentalismo é o primeiro e mais básico princípio pragmatista. Este princípio pode ser compreendido como a aversão a qualquer modelo fixado de realidade, qualquer universalismo nos usos da racionalidade. A verdade platônica contida em uma essência é terminantemente descartada em prol do olhar sistematizado para a complexidade de universos co-existentes. Há, entre estes, fronteiras sensíveis, maleáveis, concorrentes, invariavelmente abertas ao diálogo, ao intercâmbio de experiências e descobertas que articula a sobrevivência das sociedades. (JACOB, ALBRECHSEN, 1999).

O segundo elemento fundamental do pragmatismo é o *fallibilism*, ou a “faculdade da falha”. Este amplia o argumento pragmático contra a dominação de modelos rígidos e formalizados de realidade. Isto significa pensar que toda interpretação da experiência é uma tentativa de apreensão do mundo e está aberta a falhas, erros, ou seja, é sempre passível de re-análises dentro de contextos competitivos, colocando em evidência a importância da preservação de dados de pesquisa (SAYÃO, SALES, 2013) bem como dos resultados negativos (SAYÃO SALES, FELIPE, 2020) para o avanço da ciência. Paralelo a esta categoria caminha outro elemento, a contingência. A noção pragmática de contingência não colocará em evidência apenas a instabilidade do universo em que os indivíduos operam, mas também o papel penetrante e inescapável da possibilidade do imprevisto, da incerteza, do inesperado – e, principalmente, ilumina o elemento da contextualidade, ou seja, cada evento tem seu contexto específico. (JACOB, ALBRECHSEN, 1999).

Uma vez imersos em universos fragmentados, com suas experiências peculiares de identificação e reivindicação, os indivíduos devem ser estudados a partir de suas práticas sociais compartilhadas. Isto quer dizer, o pragmatismo volta-se para a socialidade das comunidades que se desenvolvem nos interditos do cotidiano, e enxerga que, para superar o relativismo da contingência em suas investigações, deve ir ao encontro de tais comunidades. É ali, no estudo das relações sociais e suas formas de interpretação que o pragmatismo acredita ser possível compreender os estratos mínimos de uma tradição. Para isso, a linguagem, ou seja, a tessitura cultural do conhecimento, é objeto fundamental de análise. (JACOB, ALBRECHSEN, 1999).

O último elemento, o pluralismo, reafirma a visão humanista de uma proposta pragmática. A partir do pensamento plural a corrente pragmática entende a constituição do mundo como algo formado pela multiplicidade de tradições, perspectivas ideológicas, utopias, instituições políticas e organizações sociais. O pluralismo pode, em uma visão superficial, conduzir ao relativismo acima citado. Contra isso, o olhar pragmático fundamenta uma análise crítica a partir das três categorias iniciais de sua estrutura, ou seja, na formação do diálogo entre teorias divergentes procura identificar as tentativas de um fundamentalismo, as possibilidades de incerteza e a dinâmica da contingência. Fronteiras entre as teorias divergentes começam a ser derrubadas a partir da confrontação e da argumentação. (JACOB, ALBRECHSEN, 1999).

A quebra de tais fronteiras (o dualismo essencialista) só pode vir – e aqui existe a pragmática crítica – com a criação permanente da plataforma de disposição da comensurabilidade e da hermenêutica, isto é, o espaço de deliberação para os sujeitos falarem, ouvirem e avaliarem as consequências de seus projetos de intervenção científica. No plano da Organização do Conhecimento, trata-se no enfoque da



pluralidade dos saberes, a busca por uma visão antropológica da informação, que permite um ponto de vista de integração pragmática de dados para cada domínio em sua dinâmica de transformação através da apresentabilidade da linguagem.

- Base teórico-metodológica: operacionalizando a organização do conhecimento pela via wittgensteiniana

Como a perspectiva do presente trabalho parte de um estudo que se concentra no papel da linguagem no contexto das linguagens documentárias para inovação, retomamos, a partir da experiência histórica do *Ecce Liber* e seus projetos em Ciência da Informação e Organização do Conhecimento, um método baseado na Filosofia da Linguagem. A leitura para a construção do método tem no filósofo austríaco Ludwig Wittgenstein o aporte principal. O método específico selecionado, retirado do pensamento do filósofo, é a “apresentação panorâmica”.

A operacionalização das etapas de coleta da pesquisa se dá a partir de três perspectivas metodológicas: bibliográfica-documental, teórico-institucional e desenvolvimento de redes. O passo inicial do projeto se dá a partir da identificação, do cotejamento e do estudo de fontes bibliográficas referentes ao domínio, partindo de experiências já constituídas. A segunda etapa procura correlacionar o papel da linguagem e das instituições informacionais (bibliotecas, arquivos e museus), com vistas a identificação de centros de coleta, processamento e memória de dados. A terceira etapa visa a integração pragmática de dados em um portal na web dedicado à política de visibilização da produção de saberes, a partir da linguagem. Procuramos a seguir debater o método de análise dos documentos levantados apontado como referência para a construção do trabalho, além de descrevê-lo e analisá-lo a partir do objeto de pesquisa lançado.

A descrição da proposta metodológica, que nos leva do plano filosófico ao plano empírico cultural demandado pela complexidade do domínio, está estruturada, fundamentalmente, nas seguintes obras: *Investigações Filosóficas*; *Livro Azul*; *Livro Marrom*; *Observações Filosóficas*; *Da Certeza*; *Big Typescript*. Dentre os principais interlocutores na leitura do pensamento wittgensteiniano em sua abordagem pragmática, adotamos Mauro Lúcio Leitão Condé. A principal fonte direta de reflexão de nossa metodologia está contida na obra “*Investigações Filosóficas*”, que demarca a segunda fase do pensamento do filósofo, orientada para a pragmática da linguagem ordinária. É aqui que encontramos a noção de “apresentabilidade” como forma crítica de contestar a relação de “representação” no plano filosófico e social.

O motivo da seleção do método, sustentado pela filosofia da linguagem, pode ser justificado, do ponto de vista do domínio, pela capacidade crítico-filosófica da

percepção pragmática de integração dos saberes proposta por Wittgenstein. Tomando por base o pensamento da segunda fase de Wittgenstein (1979), chegamos a uma reflexão que se pauta nas seguintes construções filosóficas, que permitem uma metodologia para: a) construir o conceito de “linguagem ordinária” científica para reconhecer os ramos da árvore do conhecimento como formas de vida; b) definir a gramática como epistemologia, tomando a constituição curricular e seus elementos conceituais como “linguagem primitiva”; c) elaborar o conceito de semelhanças de família (*Familienähnlichkeit*) como indícios pragmáticos de interdisciplinaridade; e) compreender os estratos conceituais – saberes primitivos – que conduzem à formalização de disciplinas como ferramentas gramaticais e as disciplinas como “jogos de linguagem” (*Sprachspiel*).

A concepção de “jogos de linguagem” nos permite determinar alguns pressupostos clássicos da filosofia wittgensteiniana, que tomam a significação a partir do uso dos termos na linguagem – ou, simplesmente, o significado como uso. Os “jogos de linguagem” têm como referência o contexto da vida social, e neste atuam. Descrever uma linguagem significa observar uma “forma de vida”, a partir de sua “apresentação panorâmica”. O jogo de linguagem é uma parte, ou parte de uma atividade, de uma “forma de vida”. As regras da “gramática” que rege esta forma de vida são também flexíveis, e se transformam junto das necessidades apresentadas na atuação dos indivíduos em cada comunidade, ou seja, uma “forma de vida”. Será através da “gramática” que poderemos constituir critérios de racionalidade para nossa atuação. Segue-se aqui a possibilidade de compreender gramáticas específicas para inovação e visibiliza-las – “apresenta-las”.

Para Wittgenstein, compreender uma palavra, isto é, aproximar-se de seu significado, são ações que se dão dentro de um sistema de linguagem. É visualizando este sistema, onde os signos podem ser apresentados como passíveis de significação, que o filósofo de Viena reorienta sua noção de lógica. Há que se perguntar pelo uso das palavras para tecer a aproximação, que pode ser elaborada pela sua descrição. Assim, antes de um espelho, o sentido de uma pergunta é “o método de responder a essa pergunta”. O método no pensamento wittgensteiniano, voltado para uma “apresentação panorâmica” se desenvolve. Se, em um dado momento, nos questionávamos sobre o método para chegar às respostas, a filosofia da linguagem ordinária do filósofo de Viena nos apresenta um método para realizar perguntas. Perguntar pelo uso, partindo da descrição, é uma das sugestões mais importantes desta abordagem. O método pode estabelecer, ao final, o próprio sentido.

Nas Observações Filosóficas, encontramos a afirmação de um método para a filosofia que não só está em fase de construção, como é a face de uma construção

permanente. A filosofia é parte desta construção inacabada, posto que é diária. O próprio método o filósofo utilizará ao longo da década de 1930 para construir as Investigações Filosóficas posto que o método é a possibilidade filosófica já que permite aproximar os significados dos instrumentos que nos ajudam a agir no mundo, ou seja, os conceitos. Só podemos, afirma o filósofo austríaco, “indicar o comprimento de um objeto se dispor de um método para encontrar o objeto.” (WITTGENSTEIN, 2005, p. 56)

E é este complexo sistema de linguagem, munido de suas regras, aquele que constituirá os limites de nossa compreensão, posto que demarca a gramática. Somos livres, mas dentro de fronteiras de linguagem. Não podemos conhecer o mundo que há fora de nossas possibilidades de nossa língua. “A gramática proporciona à linguagem os graus necessários de liberdade” (WITTGENSTEIN, 2005, p. 57). Um dos erros da filosofia é procurar uma espécie de compreensão, segundo a leitura wittgensteiniana, que está além de nossa linguagem. Isto, em sua visão, é impossível. Uma determinada “naturalidade do mundo” está, justamente, no “fato de que a linguagem pode referir-se, e de fato se refere, somente a ele”. Isto se dá, esclarece o filósofo de Viena, porque a “linguagem só obtém sua forma de significar daquilo que significa, do mundo, portanto, não é concebível nenhuma linguagem que não represente este mundo.” (WITTGENSTEIN, 2005, p. 65). Eis o plano da “apresentabilidade”, o modo como o mundo se “apresenta” a partir da pluralidade de seus significados, da vida dinâmica, do local para o global, de cada jogo de linguagem.

- Base empírica: coleta, reunião e análise

A partir das estruturas teórico conceitual, filosófica e metodológica, a pesquisa pretende identificar, selecionar, reunir e ordenar as linguagens documentárias brasileiras, avançando para a etapa analítica a partir das seguintes variáveis:

- Caracterização da linguagem documentária (classificação, tesouro, ontologia, etc.);
- Condição técnica (tipo de suporte e de recursos de uso e de visualização)
- Usabilidade (modalidades de apropriação)
- Acesso (tendências de perfis de uso e de reuso)
- Apropriação, construção e aplicação de métodos de ensino e aprendizagem em organização do conhecimento a partir de mestrandos e doutorandos do PPGCI IBICT UFRJ.

## 5 Dos resultados esperados

As etapas da pesquisa desenvolvidas a partir de um cronograma de atividades de execução de 36 (trinta e seis) meses são distribuídas da seguinte forma: A. Estudo de natureza bibliográfica e documental de abordagens teórico-metodológicas em organização do conhecimento com vistas à inovação; B. Mapeamento da produção das linguagens documentárias no território nacional; C. Revisão do corpus documental e metadocumental identificado; D. Desenvolvimento final do corpus; E. Análise dos dados; F. Discussão dos dados; G. Relatório final; H. Produção bibliográfica.

A partir do curso das atividades os produtos esperados como resultado da execução do projeto, com previsão de cronograma de entregas anuais são:

- Primeiro ano: conclusão do estudo de natureza bibliográfica e documental abordagens teórico-metodológicas em organização do conhecimento, bem como término da revisão do corpus documental e metadocumental constituído pelas linguagens documentárias do Brasil;
- Segundo ano: conclusão das etapas de desenvolvimento do corpus linguagens documentárias brasileiras e análise das suas condições teórico-técnicas;
- Terceiro ano: apresentação dos resultados referentes ao quadro geral das linguagens documentárias, seu atual estado e suas dinâmicas de uso e de reuso, fomentando integração pragmática de dados em um portal na web dedicado à política de visibilidade da produção de saberes com vistas à inovação no país.

A partir destas etapas, espera-se como resultados a consecução do desenvolvimento de um modelo de identificação e de reunião das linguagens documentárias, com foco na contribuição para o desenvolvimento econômico, político, social e científico do Brasil, tendo como horizonte final a inovação como forma de fundamentação contínua da estabilidade econômica do território.

Dado o cenário de complexidade econômica, política, social e científica da correlação organização do conhecimento e inovação, o quadro potencial de impacto dos resultados do ponto de vista técnico-científico, de inovação, difusão, sócio-econômico e ambiental é assim sintetizado:

- Impacto para socioeconômico: ampliação do potencial de interrelação e de descoberta de insumos para o desenvolvimento;
- Impacto para o acesso à informação da produção do conhecimento por parte de pesquisadores, estudantes e da sociedade em geral;

- Impacto para políticas públicas: visibilidade de linguagens e de sistemas abertos e para tomadas de decisão;
- Impacto para a organização do conhecimento: aprofundamento do desenvolvimento de teorias e métodos em organização do conhecimento.

### **Agradecimento**

À comunidade discente das disciplinas Organização dos Domínios do Conhecimento e Sistemas de Organização do Conhecimento.

### **Financiamento**

A pesquisa foi desenvolvida a partir do fomento do Conselho Nacional para o Desenvolvimento Científico e Tecnológico (CNPq) e da Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (Faperj).

### **Referências**

ARISTÓTELES. **Órganon: Categorias, Da interpretação, Analíticos anteriores, Analíticos posteriores, Tópicos, Refutações sofisticas**. 2. ed. Bauru, SP: Edipro, 2010.

ARISTÓTELES. **Metafísica**: ensaio introdutório, texto em grego com tradução e comentário de Giovanni Reale. São Paulo: Loyola, 2005

ARISTÓTELES. **Rhétorique**. Paris: Gallimard, 1991.

ARISTÓTELES. **Poética**. Porto Alegre: Editora Globo, 1966.

BERGER, P. L.; LUCKMAN, T. **A construção social da realidade**: tratado de sociologia do conhecimento. Tradução Floriano de Souza Fernandes. 25<sup>a</sup> ed. São Paulo: Vozes, 2005.

BAKHTIN, Mikhail. **Marxismo e filosofia da linguagem**. São Paulo: Hucitec, 2012.

BAKHTIN, Mikhail. **A cultura popular na Idade Média e no Renascimento**: o contexto de François Rabelais. São Paulo: Hucitec, 2010.

BRIER, Soren. **Cybersemiotics**: a new interdisciplinary development applied to the problems of knowledge organization and document retrieval in information science. *Journal of Documentation*, v. 52, n. 3, p. 296-344, sep. 1996.

CAMPOS, Maria Luiza de A.; GOMES, Hagar E. Organização dos domínios do conhecimento e os princípios ranganhathianos. **Perspect. Cienc. Inf.**, Belo Horizonte, v. 8, n. 2, p. 150-163, jul./dez. 2003.

CAPURRO, R. What is Information Science for? a philosophical reflection In: VAKKARI, P.; CRONIN, B. (Eds.). **Conceptions of Library and Information Science**; historical, empirical and theoretical perspectives. In: INTERNATIONAL CONFERENCE FOR THE CELEBRATION OF 20TH ANNIVERSARY OF THE DEPARTMENT OF INFORMATION STUDIES, UNIVERSITY OF TAMPERE, FINLAND. 1991. Proceedings... London, Los Angeles: Taylor Graham, 1992. p.82-96.

CASSIRER, Ernst. **A filosofia das formas simbólicas**; 3a parte: Fenomenologia do conhecimento. São Paulo: Martins Fontes, 2011.

CASSIRER, Ernst. **A filosofia das formas simbólicas**; 2a parte: O pensamento mítico. São Paulo: Martins Fontes, 2004.

CASSIRER, Ernst. **A filosofia das formas simbólicas**; 1a parte: A linguagem. São Paulo: Martins Fontes, 2001.

CASSIRER, Ernst. **Ensaio sobre o homem**: introdução a uma filosofia da cultura humana. São Paulo: Martins Fontes, 1994.

COELHO-SILVA, Jacionira. **Arqueologia do Médio São Francisco**. Indígenas, vaqueiros e missionários. Tese de Doutorado. UFPE/PPGH. Recife, 2003.

DAY, Ronald. **Poststructuralism and information studies**. Annual review of information science social and technology (ARIST), v. 39, p. 575-609, 2005.

ECO, Umberto. **A vertigem das listas**. Rio de Janeiro: Record, 2010.

ECO, Umberto. **Semiótica e filosofia da linguagem**. Lisboa: Instituto Piaget, 1984.

FROHMANN, Bernd. The power of images: a discourse of images: a discourse analysis of cognitive viewpoint. **Journal of Documentation**, v. 48, n. 4, p. 365-386, 1992.

GARCÍA GUTIÉRREZ, A. Desclassification in Knowledge Organization: a postepistemological essay. **Transinformação**, Campinas, v.23, n.1, p. 5-14, 2011.

GIANNOTTI, J. A. **Apresentação do mundo**: considerações sobre o pensamento de Ludwig Wittgenstein. São Paulo: Cia das Letras, 1995.

GONZÁLEZ DE GÓMEZ, Maria Néida. A representação do conhecimento e o conhecimento da representação: algumas questões epistemológicas. **Ci. Inf.**,

Brasília, v. 22, n. 3, p. 217-222, set./dez. 1993;

GONZÁLEZ DE GÓMEZ, Maria Nélide. Comentários ao artigo “Hacia um nuevo paradigma em bibliotecologia”. **Transinformação**, Campinas, v. 8, n. 3, p. 44-56, set./dez. 1996.

GONZÁLEZ DE GÓMEZ, Maria Nélide. Da organização do conhecimento às políticas de informação. **Informare**: caderno do programa de pós-graduação em ciência da informação, Rio de Janeiro, v. 2, n. 2, p. 58-66, jul./dez. 1996.

GUIMARÃES, José A. C.; PINHO, Fabio; MILANI, Suellen. Theoretical Dialogs About Ethical Issues in Knowledge Organization: García Gutiérrez, Hudon, Beghtol, and Olson. **Knowl. Org.**, v. 43, n. 5, 2016.

JOSEPH, Miriam. **O Trivium**: as artes liberais da lógica, gramática e retórica: entendendo a natureza e a função da linguagem. São Paulo: É Realizações, 2008.

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, jul. 1995.

HJORLAND, B. Domain analysis in information science: eleven approaches traditional as well as innovative. **Journal of Documentation**, Londres, v. 58, n. 4, p. 422-462, 2002.

JACOB, E.K.; ALBRECHTSEN, H. **When essence becomes function**: Post-structuralist implications for an ecological theory of organisational classification systems. In T.D. Wilson & D.K. Allen. Exploring the contexts of information behaviour: proceedings of the second international conference on research in information needs, seeking and use in different contexts. London: Taylor Graham, 1999. p. 519-534.

MENDES et. al. A organização das mulheres extrativistas na região Sudoeste mato-grossense, Brasil. **Estudos Feministas**, Florianópolis, v. 22, n.1, p. 71-89, jan.-abr. 2014.

MARTÍNEZ ALIER, Joan. **O Ecologismo dos Pobres**: conflitos ambientais e linguagens de valoração. São Paulo: Contexto, 2007.

OCDE. **Manual de Oslo**: proposta de diretrizes para coleta e interpretação de dados sobre inovação tecnológica. Brasília: FINEP, 1997.

OLSON, Hope. **Exclusivity, Teleology and Hierarchy**: Our Aristotelean Legacy.

Know!. Org., v. 26, n. 2, p. 65-73, 1999.

RENDÓN ROJAS, Miguel A. R. Hacia um nuevo paradigma em bibliotecologia. **Transinformação**, Campinas, v. 8, n. 3, p. 17-31, set./dez. 1996.

SALDANHA, Gustavo S. Sobre a O<sup>2</sup>S<sup>2</sup>O, de tesouro à Bourdieu: linguagem simbólica e a organização ordinária dos saberes socialmente oprimidos In.: LUCAS, Elaine R. de O.; SILVEIRA, Murilo A. A. da. **A Ciência da informação encontra Pierre Bourdieu**. Recife: Ed. Universitária da UFPE, 2017. p. 173-196.

SALDANHA, Gustavo F. **Linguagens documentárias fluminenses para a inovação: organização do conhecimento para ciência, cultura e sociedade no Estado do Rio de Janeiro**. Rio de Janeiro: 2021. Projeto de pesquisa sob financiamento do Edital Jovem Cientista do Estado (2019-2021) - Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (Faperj).

SAYÃO, Luis Fernando, SALES, Luana Farias. Curadoria digital: um novo patamar para preservação de dados digitais de pesquisa. **Informação & Sociedade: Estudos**, 2013.

SAYÃO, Luis Fernando; SALES, Luana Farias; FELIPE, Carla Beatriz Marques. A ciência invisível: a publicação dos resultados negativos de pesquisa. **Transinformação**, v. 33, 2021.

TESAURO, Emanuelle. **Il cannocchiale aristotelico**. Turim: Bartolomeo Zauatta, 1670. TORRES, Maurício. O escriba e o narrador: A memória e a luta pela terra dos ribeirinhos do Alto Tapajós. **Tempo Social**, revista de sociologia da USP, v. 26, n. 1, p. 233-257, 2014.

WITTGENSTEIN, L. **Tratado Lógico Filosófico; Investigações filosóficas**. Tradução M. S. Lourenço. 3ª ed. Lisboa: Fundação Calouste Gulbenkian, 2002;

WITTGENSTEIN, L. **Investigações Filosóficas**. 20 ed. Tradução de José Carlos Bruni. São Paulo: Abril Cultural, 1979. (Os Pensadores);

WITTGENSTEIN, L. **O livro azul**. Lisboa: Ed.70, 1992a. 125 p.

WITTGENSTEIN, L. **O livro castanho**. Rio de Janeiro: Edições 70, 1992b.

WITTGENSTEIN, L. **Observações filosóficas**. Tradução de Adail Sobral e Maria Stela Gonçalves. São Paulo: Loyola, 2005.

VIGNAUX, Georges. **O demónio da classificação: pensar, organizar**. Lisboa: Instituto Piaget, 2000.



# **A organização do conhecimento na Fiocruz: uma investigação dos seus sistemas de organização do conhecimento**

Denizard Costa da Silva<sup>1</sup>, Elaine Lucia da Silva<sup>2</sup>, Karina Veras  
Praxedes dos Santos<sup>3</sup> e Thaís de Oliveira Sant'Anna Campanhã<sup>4</sup>

---

## **1 Introdução**

ESTE ESTUDO APRESENTA O RESULTADO DO MAPEAMENTO DOS SISTEMAS DE Organização do Conhecimento (SOC) da Rede de Bibliotecas da Fundação Oswaldo Cruz (Fiocruz), os quais abrangem todos os tipos de esquemas que organizam e representam o conhecimento, por exemplo, as classificações, lista de termos, taxonomias, tesouros, ontologias, entre outros.

Desde sua origem em 1900, como Instituto Soroterápico Municipal do Rio de Janeiro, a Fiocruz estabeleceu-se como uma organização que atua em pesquisa, educação, desenvolvimento tecnológico e inovação em saúde, e para tal gera e mantém acervos de conhecimento valiosos em seus sistemas de informação. É a principal protagonista da Ciência em Saúde no Brasil. Dada a sua grande extensão em unidades técnico-científicas, possui diversas Unidades de Informação que tratam de assuntos correlatos ou multidisciplinares à Ciência em Saúde.

É importante enfatizar que uma instituição que tem como parte de sua missão promover a saúde e o desenvolvimento social, gerar e difundir conhecimentos científicos e tecnológicos, deve ter como premissa básica um esforço contínuo para

---

1 Mestrando em Biblioteconomia, Universidade Federal do Estado do Rio de Janeiro, denizcosta@gmail.com.

2 Doutoranda em Ciência da Informação IBICT/UF RJ, Mestra em Saúde Pública, Fundação Oswaldo Cruz, elaineluciarh@gmail.com.

3 Doutoranda em Ciência da Informação IBICT/UF RJ, Mestra em Arquivologia, Fundação Oswaldo Cruz, karinakvps@yahoo.com.br

4 Mestranda em Biblioteconomia, Universidade Federal do Estado do Rio de Janeiro, thaisantanna@hotmail.com.

uma comunicação eficaz na organização e eficiente recuperação da informação em suas bibliotecas.

Neste sentido, Dahlberg (2006) reforça que a organização do conhecimento é a ciência que ordena a estruturação e sistematização dos conceitos, de acordo com suas características. A partir dessa organização do conhecimento criam-se instrumentos que apresentam a interpretação organizada e estruturada do objeto, chamados de Sistemas de Organização do Conhecimento.

Na área da Ciência da Informação, os Sistemas de Organização do Conhecimento

são sistemas conceituais semanticamente estruturados que contemplam termos, definições, relacionamentos e propriedades dos conceitos. Na organização e recuperação da informação, os SOC cumprem o objetivo de padronização terminológica para facilitar e orientar a indexação e os usuários (CARLAN; MEDEIROS, 2011, p. 54).

Segundo Moraes e Moreira (2018), a qualidade da recuperação da informação em um sistema de informação documentária depende dos procedimentos e ferramentas utilizados no seu processo de organização e representação. Assim, torna-se relevante compreender como a Fundação Oswaldo Cruz tem uma base semântica que dê apoio à busca e recuperação da informação pelo usuário e pelos profissionais das bibliotecas.

Portanto, o objetivo geral deste trabalho é investigar os SOC utilizados nas Unidades de Informação da Fiocruz, caracterizando assim, a Rede de Bibliotecas da Fiocruz como nosso campo empírico.

Traçamos como nossos objetivos específicos:

- a) mapear as Unidades de Informação existentes na Fiocruz;
- b) identificar os SOC apresentados por cada unidade como utilizados nos seus processos de organização do conhecimento; e
- c) conhecer as dificuldades/facilidades que estes possam oferecer.

Oportunamente, buscaram-se informações sobre possíveis projetos em planejamento ou implantação para desenvolver melhorias ou novos SOC para suprir alguma dificuldade conhecida.

## **2 Procedimentos metodológicos**

A fim de atingir os objetivos propostos, esta pesquisa se caracteriza como exploratória e quali-quantitativa (MINAYO, 1992). Com o intuito de descrever a com-

plexidade do problema como um todo e no contexto ao qual se aplica, necessitou-se de um estudo aprofundado diante de sua complexidade.

As técnicas de coleta de dados no trabalho foram: pesquisa bibliográfica, busca nos *sites* da Fiocruz e entrevistas *on-line* com representantes das bibliotecas, com a finalidade de compreender alguns aspectos relacionados à pesquisa, ao cenário e aos Sistemas de Organização do Conhecimento utilizados pelas Bibliotecas da Rede da Fiocruz.

Os contatos e entrevistas foram realizados remotamente, utilizando as tecnologias disponíveis, em virtude do isolamento social provocado pela pandemia do Coronavírus. O recurso de videoconferência viabilizou a execução em função do contexto a que todos, pessoas e instituições, estiveram submetidos.

Para alcançar o primeiro objetivo específico de mapear as Unidades de Informação existentes na Fiocruz, recorreu-se a uma exploração via *browser* nos diversos sítios das unidades da Fundação, notadamente apontados sob o domínio *fiocruz.br*.

Pôde-se então, elaborar o primeiro olhar sobre a diversidade de Unidades de Informação existentes na Fundação, considerando-se, inicialmente, as unidades físicas e, em seguida, as contextualizadas como recurso virtual.

Seja durante o “*brausear*” ou em decorrência das informações recebidas durante a etapa de entrevistas, foi possível enumerar algumas das unidades institucionais que são diretamente atendidas ou vinculadas às Unidades de Informação mapeadas.

Como Unidades de Informação física, foram enumeradas no Quadro 1 as seguintes unidades:

**Quadro 1 - Mapeamento de Unidades de Informação físicas**

<b>Unidade de Informação<sup>5</sup></b>	<b>Unidades Técnico-Científica Atendidas / Vinculadas<sup>6</sup></b>
<b>Biblioteca da Fiocruz Brasília</b>	Diretoria Regional de Brasília (DIREB); Escola Fiocruz de Governo
<b>Biblioteca da Saúde da Mulher e da Criança</b>	Instituto Fernandes Figueira (IFF); Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT)
<b>Biblioteca de Ciências Biomédicas Eurydice Pires de Sant'Anna (BEPS)</b>	Instituto Gonçalo Moniz (Fiocruz Bahia)
<b>Biblioteca de Ciências da Saúde Prof. Zigman Brener</b>	Instituto René Rachou (Fiocruz Minas)
<b>Biblioteca de Educação e Divulgação Científica Ilone Seibel</b>	Museu da Vida; Casa de Oswaldo Cruz (COC)

5 Destaca-se (fundo cinza) as unidades com as quais obteve-se sucesso de agenda para entrevista.

6 Destaca-se (fundo cinza) as unidades logradas no município do Rio de Janeiro.

**Quadro 1 - Mapeamento de Unidades de Informação físicas**

<b>Unidade de Informação<sup>5</sup></b>	<b>Unidades Técnico-Científica Atendidas / Vinculadas<sup>6</sup></b>
<b>Biblioteca de História das Ciências e da Saúde</b>	Casa de Oswaldo Cruz (COC)
<b>Biblioteca de Manguinhos</b>	Instituto Oswaldo Cruz (IOC); Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT); Instituto Nacional de Infectologia Evandro Chagas (INI); Instituto de Ciência e Tecnologia em Biomodelos (ICTB); Instituto de Tecnologia em Imunobiológicos (Bio-Manguinhos)
<b>Biblioteca de Medicamentos e Fitomedicamentos</b>	Instituto de Tecnologia em Fármacos (Farmanguinhos)
<b>Biblioteca de Saúde Pública</b>	Escola Nacional de Saúde Pública Sergio Arouca (ENSP); Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT);
<b>Biblioteca do CESTEHE</b>	Centro de Estudos da Saúde do Trabalhador e Ecologia Humana (CESTEHE/ ENSP)
<b>Biblioteca do INCQS</b>	Instituto Nacional de Controle de Qualidade em Saúde (INCQS)
<b>Biblioteca do ICC</b>	Instituto Carlos Chagas (ICC/Fiocruz Paraná)
<b>Biblioteca do Núcleo de Informação e Documentação (NIC/CLAVES)</b>	Centro Latino-Americano de Estudos de Violência e Saúde Jorge Careli (CLAVES) - integrante da Escola Nacional de Saúde Pública (ENSP) e do Instituto Fernandes Figueira (IFF)
<b>Biblioteca Dr. Antônio Levino</b>	Instituto Leonidas e Maria Deane (Fiocruz Amazonas)
<b>Biblioteca Emília Bustamante (BEB)</b>	Escola Politécnica de Saúde Joaquim Venâncio (EPSJV)
<b>Biblioteca Instituto Aggeu Magalhães (IAM)</b>	Instituto Aggeu Magalhães (Fiocruz Pernambuco)
<b>Biblioteca Livre Palácio Itaboraí</b>	Fórum Itaboraí: Política Ciência e Cultura na Saúde (Petrópolis)
<b>Biblioteca Walter Mendes</b>	Centro de Referência Professor Hélio Fraga (CRPHF/ENSP)
<b>Centro de Documentação da GESTEC (CDG)</b>	Coordenação de Gestão Tecnológica da Fiocruz

Fonte: Os autores (2021).

Ainda com o propósito de cumprir o objetivo de mapeamento, também julgou-se oportuno entrevistar representantes de Unidades de Informação Virtuais. O Quadro 2 foi elaborado para apresentar as Unidades de Informação Virtuais.

**Quadro 2 - Mapeamento de Unidades de Informação virtuais**

<b>Unidade de Informação<sup>7</sup></b>	<b>Unidades Técnico-Científica Atendidas / Vinculadas</b>
<b>Temáticas</b>	
BVS Aleitamento Materno	Biblioteca da Saúde da Mulher e da Criança
BVS Bioética e Diplomacia da Saúde	Núcleo de Estudos sobre Bioética e Diplomacia em Saúde (NETHIS)
BVS Carga de Doença	Núcleo de Estudos de Carga Global de Doença
BVS Determinantes Sociais em Saúde	Centro de Estudos, Políticas e Informação sobre Determinantes Sociais da Saúde da Escola Nacional de Saúde Pública (CEPI-DSS-ENSP)
BVS Doenças Infecciosas e Parasitárias	Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT)
BVS Educação Profissional em Saúde	Escola Politécnica de Saúde Joaquim Venâncio (EPSJV)

<sup>7</sup> Destaca-se (fundo cinza) as unidades com as quais agendou-se entrevista.

**Quadro 2 - Mapeamento de Unidades de Informação virtuais**

<b>Unidade de Informação<sup>7</sup></b>	<b>Unidades Técnico-Científica Atendidas / Vinculadas</b>
BVS Envelhecimento	Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT)
BVS História da Poliomielite	Casa de Oswaldo Cruz (COC)
BVS História e Patrimônio Cultural da Saúde	Casa de Oswaldo Cruz (COC)
BVS Integralidade em Saúde	Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT)
BVS Pensamento Social	Casa de Oswaldo Cruz (COC)
BVS Saúde dos Povos Indígenas	Escola Nacional de Saúde Pública Sergio Arouca (ENSP)
BVS Saúde Pública	Biblioteca de Saúde Pública; Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT)
BVS Violência e Saúde	Departamento de Estudos sobre Violência e Saúde Jorge Careli (CLAVES)
Biográficas	
BV Oswaldo Cruz	Casa de Oswaldo Cruz (COC)
BV Sergio Arouca	Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT)
BVS Adolpho Lutz	Casa de Oswaldo Cruz (COC)
BVS Carlos Chagas	Casa de Oswaldo Cruz (COC)
Repositórios	
Arca - Repositório Institucional da Fiocruz <sup>8</sup>	Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT)
Base Teses Fiocruz <sup>9</sup>	Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT)
Catálogo Mourisco <sup>10</sup>	Rede de Bibliotecas da Fiocruz
Portal de Periódicos Fiocruz <sup>11</sup>	Fundação Oswaldo Cruz/Presidência

Fonte: Os autores (2021).

Assim, chegamos a um número de 19 unidades físicas e 22 unidades virtuais. Dada a extensa lista de unidades da instituição que foram mapeadas, foi necessário definir algum critério para reduzir a atividade a fim de que fosse possível elaborar uma agenda exequível para a realização das entrevistas.

Considerando o contexto pandêmico sobre o qual o país encontrava-se durante a realização desta pesquisa, adotou-se como critério, centralizar os esforços de contato com representantes das unidades físicas que são logradas no município do Rio de Janeiro.

8 Consultas a partir de 2018 (FUNDAÇÃO OSWALDO CRUZ, [2018?]).

9 Descontinuada em 02/2018, mantida no ar para consulta do intervalo 1996 a 2018 (FUNDAÇÃO OSWALDO CRUZ, [2018?]).

10 Trata-se do software ALEPH implementado para promover a reunião de mais de 30 mil textos completos a partir da colaboração das bibliotecas integrantes da Rede de Bibliotecas Fiocruz.

11 Plataforma de hospedagem dos periódicos da Fiocruz. Disponível em <https://periodicos.fiocruz.br/>.

Atribuído o critério de corte, deu-se início ao esforço de contato, via telefone e *e-mail*, requerendo oportunidade de entrevista por videoconferência. Obteve-se sucesso para uma entrevista com a representante da Rede de Bibliotecas, o que auxiliou a definição de um segundo critério de corte, no qual dado que a Rede é, por si, uma virtualização das unidades físicas somadas a outras unidades nativas virtuais, decidiu-se eleger uma unidade de Biblioteca Virtual para considerar em nosso estudo o contexto das Bibliotecas Virtuais.

Também selecionamos o repositório institucional ARCA como necessário para determos nosso olhar. Consideramos assim que, dado o contexto social mencionado, tenhamos representado as unidades virtuais em nosso estudo, de maneira satisfatória.

Emitidos os contatos, obteve-se um sucesso de sete aceites para concessão de entrevista, junto às dez bibliotecas físicas logradas no município do Rio de Janeiro (correspondendo a uma aceitação de 70%). À estas sete, incluiu-se a Biblioteca Virtual Oswaldo Cruz, de caráter biográfico, o repositório Arca desenvolvido sobre a plataforma DSpace, e a própria Rede de Bibliotecas, representada por sua coordenação. Perfazendo uma lista de 10 entrevistas realizadas:

- Rede de Bibliotecas da Fiocruz
- Biblioteca da Saúde da Mulher e da Criança
- Biblioteca de Educação e Divulgação Científica Ilone Seibel
- Biblioteca de História das Ciências e da Saúde
- Biblioteca de Manguinhos
- Biblioteca de Medicamentos e Fitomedicamentos
- Biblioteca de Saúde Pública
- Biblioteca Emília Bustamante (BEB)
- BV Oswaldo Cruz
- Arca - Repositório Institucional

Para atingir o segundo e terceiro objetivos específicos, optou-se pela coleta de dados por meio de entrevistas de caráter exploratório, na qual se permite “eventuais indagações ou levantamento de dados e informações que não estejam contempladas no formulário” (KAUARK; MANHÃES; MEDEIROS, 2010, p. 62). Assim, os profissionais entrevistados ficaram à vontade para trazer questões relevantes, mesmo que não contempladas no roteiro inicial. Este instrumento foi escolhido, pois, diferente do questionário, permite adentrar as subjetividades que possam vir a ocorrer na prática de trabalho e de uso dos Sistemas de Organização do Conhecimento.

Encerradas as etapas de mapeamento das Unidades de Informação e contato com os representantes das mesmas, iniciou-se a etapa das entrevistas virtuais, que foram realizadas entre os meses de setembro e outubro de 2020, via plataforma *Google Meet*.

Finalizadas as entrevistas, foram iniciadas as etapas de análises, o que nos proporcionou aprofundamento e consolidação das considerações finais.

### **3 Os Sistemas de Organização do Conhecimento na Fiocruz**

Mudanças ao longo do tempo provocaram mudanças na Organização do Conhecimento. A invenção da escrita e posteriormente da imprensa, a revolução industrial e as novas tecnologias, somados ao grande número de informações trouxeram consequências para a área. Oferecer acesso é o principal objetivo de um centro de informação. Assim sendo, a classificação e a indexação são essenciais para a recuperação da informação, e os instrumentos para realizar essas tarefas facilitam a vida dos usuários e dos profissionais.

Os Sistemas de Organização do Conhecimento são instrumentos que auxiliam nessa organização, muitas vezes dentro de um contexto particularizado, como por exemplo, esquemas de classificações, ontologias, tesouros, etc. Segundo Carlan e Me-deiros (2011), as estruturas variam de esquemas simples a multidimensionais, tentando eliminar ambiguidades conceituais, e controlar sinônimos ou equivalências.

Os SOC

desempenham a função de subsidiar os processos de organização e recuperação da informação com a padronização dos termos empregados para representar os conceitos, delimitação do significado a ser interpretado e indicação precisa dos relacionamentos semânticos entre conceitos (MELO; BRASCHER, 2014).

Tabelas de classificação como a Classificação Decimal de Dewey (CDD), por exemplo, são organizadas de forma sistemática e visam, principalmente, a localização física das obras. Instrumentos como vocabulários controlados, tesouros, listas de cabeçalhos de assuntos, etc, visam, principalmente, organizar os assuntos tratados nos documentos. Todos são tipos de SOC, portanto, propõem-se a organizar o conhecimento e, assim, reunir os documentos por suas características comuns.

Considerando a grande variedade de estruturas existentes, um SOC pode ser mais ou menos apropriado no contexto de atuação de uma unidade informacional. A seleção dos SOC a serem utilizados para organizar as informações e conhecimentos em quaisquer formatos, físico ou digital, deve levar em consideração não só

sua natureza, mas também as práticas de pesquisa de seus usuários e, sempre que viável, considerar o fluxo de suas cognições.

Neste sentido, as Unidades de Informação da Fiocruz buscaram formas de se adaptar para cumprir suas demandas, a fim de oferecer um serviço eficaz de busca e recuperação, com vistas ao acesso para o usuário.

Até o ano de 2020, cada biblioteca da Fiocruz tinha seu próprio sistema de gerenciamento da informação. Mas, atualmente, a Fiocruz adota o Sistema Aleph para toda a Rede de bibliotecas, como forma de facilitar a integração entre as unidades. Os soc identificados na Fiocruz como os mais usados pelas Unidades de Informação foram: Descritores em Ciências da Saúde (DeCS) e Classificação Decimal de Dewey (CDD). O DeCS tem caráter mandatório na instituição e é utilizado por toda Rede para assegurar a interoperabilidade entre as unidades e outras bibliotecas da área de saúde da América Latina e Caribe, em especial as Bibliotecas Virtuais em Saúde (BVS).

No entanto, as unidades têm liberdade para adotar medidas que supram suas necessidades particulares, pois a Rede abarca diferentes temáticas e atende grande diversidade de usuários. Para tanto, as unidades recorrem a outros soc, seja de elaboração própria ou não, que ajudam na organização da informação e sua consequente recuperação.

O resultado das entrevistas proporcionou conhecer os soc utilizados pelas unidades conforme o Quadro 3 elaborado com o propósito de sintetizar os resultados aferidos.

**Quadro 3 - SOC utilizados nas Unidades de Informação.**

<b>Unidade de Informação</b>	<b>SOC Identificados (mesmo se não em uso)</b>
Rede de Bibliotecas da Fiocruz	DeCS, Biblioteca Nacional, Tesouro de Manguinhos (Teman) e, como orientação adicional para complementação: Library of Congress, British Library e INEP.
Biblioteca da Saúde da Mulher e da Criança	Classificação Decimal de Dewey (CDD); DeCS
Biblioteca de Educação e Divulgação Científica Iloni Seibel	CDD; DeCS; Tabela Cutter; e Lista de Áreas Temáticas da Coleção Infantojuvenil
Biblioteca de História das Ciências e da Saúde	CDD; DeCS; Tabela Cutter; Descritores Locais.
Biblioteca de Manguinhos	CDD; DeCS; Teman
Biblioteca de Medicamentos e Fitomedicamentos	CDD; DeCS; Tabela Cutter; Vocabulário controlado próprio que remete ao DeCS
Biblioteca de Saúde Pública	CDD; DeCS
Biblioteca Emília Bustamante (BEB)	CDD; DeCS; Thesaurus Brasileiro da Educação do INEP
Biblioteca Virtual Oswaldo Cruz (BVOC)	DeCS e Vocabulário Controlado da COC
Arca - Repositório Institucional	"Tabela de equivalência" elaborada para implementar a ferramenta "visualização de dados"; Indexação herdada com DeCS (por base) e demais descritores herdados da prática contextual de cada unidade alimentadora.

Fonte: Os autores (2021).



A seguir, apresentam-se as considerações obtidas em cada Unidade de Informação, físicas e virtuais, a partir das entrevistas realizadas remotamente. As informações correspondem à transcrição das falas dos profissionais entrevistados que representaram suas Unidades de Informação. Para preservar a privacidade dos profissionais, seus nomes pessoais foram omitidos, sendo indicado somente o nome da Unidade representada.

### **I. Rede de Bibliotecas da Fundação Oswaldo Cruz**

Tendo em vista o campo empírico escolhido para a realização desta pesquisa, julgou-se oportuno compreender o contexto de atuação da Rede de Bibliotecas da Fiocruz, para que, a partir dessa visão geral, fosse possível conhecer com maior esclarecimento as especificidades de cada uma das bibliotecas que, adiante, tiveram seus representantes entrevistados.

Até 2019, cada biblioteca da Fiocruz tinha seu próprio sistema de gestão da informação, porém, em 2020, houve a unificação do sistema utilizado por todas as bibliotecas da Rede, que passaram a utilizar o Aleph, nomeado na instituição como Catálogo Mourisco.

De acordo com a entrevista, um dos Sistemas de Organização do Conhecimento utilizado na Fiocruz é o vocabulário controlado DeCS (Descritores em Ciências da Saúde) e a determinação de seu emprego varia a partir do sistema de informação utilizado na instituição. Isso pode ser visualizado, por exemplo, nas Bibliotecas Virtuais em Saúde (BVS) e na Literatura Científica e Técnica da América Latina e Caribe (Lilacs) onde a utilização de um descritor DeCS é mandatória para permitir a indexação, embora possibilite o uso de palavras “soltas”, enquanto palavras-chave. Tal condição difere do que ocorre no sistema Aleph que é flexível quanto ao uso, ou não, do DeCS para a indexação, ou seja, o uso do descritor é livre, contudo a orientação é que sua escolha se dê a partir do uso de um tesouro. Para nortear essa decisão do descritor a ser usado, foi inserido no Catálogo Mourisco o Manual de entrada de dados, dividido por áreas, que recomenda as fontes que podem ser utilizadas para apoiar a indexação, a saber: DeCS, Biblioteca Nacional, Tesouro de Manguinhos (Teman) e, como orientação adicional para complementação: Library of Congress, British Library e INEP.

Ressalta-se que, a partir da integração entre as bibliotecas, foi criado, no âmbito da Rede, o Grupo de Trabalho de Indexação que está mapeando outros vocabulários controlados que já apoiavam a indexação de termos específicos nos sistemas de gestão de acervos anteriormente utilizados, portanto, serão acrescentadas ao Manual de entrada de dados. Esse mesmo grupo de trabalho está desenvolvendo a Política de Indexação e, em seguida, realizará a atualização do Teman com objetivo de normalizar e padronizar a parte da semântica que é específica da Fiocruz.

Um aspecto importante destacado é quanto à obrigatoriedade de uso do DeCS como um desafio e uma solução, ao mesmo tempo. Desafiante à medida que este SOC não abarca a diversidade das especificidades da Fiocruz, o que pode acarretar que o descritor escolhido não represente adequadamente o conteúdo ao qual se relaciona. Outro fator é que a inserção de novos descritores também não é feita de maneira muito ágil, o que atrasa algumas indexações, tal fato ocorreu com a COVID-19. Todavia, ao se pensar em qualidade para a interoperabilidade semântica entre os vários sistemas de instituições do campo da saúde, tal condição se torna necessária.

Para a classificação nas bibliotecas da Rede utiliza-se o CDD. O CDU é utilizado em apenas 1 delas: na Biblioteca do Instituto Ageu Magalhães.

Cabe destacar que para indexação de dados de pesquisa são usados diretórios que utilizam ontologias e vocabulários controlados específicos para esses objetos, principalmente para propiciar uma orientação ao pesquisador que vai admitir os dados de sua pesquisa em repositórios externos à Fiocruz, por exigência de revistas científicas, por exemplo.

Quanto ao desempenho na recuperação da informação com a utilização desses SOC foi ressaltado que o retorno não está atrelado exclusivamente ao tesouro ou ao vocabulário controlado, e sim ao sistema como um todo, considerando também neste contexto os sistemas eletrônicos utilizados para indexação. Especificamente sobre o DeCS, sua utilização nas bases próprias da BVS e Lilacs, propicia a recuperação da informação de forma trilingue: português, espanhol e inglês.

Segundo falado na entrevista, é necessária a utilização de sistemas mais modernos, com ontologias robustas que não pense mais num termo isolado, mas nas suas tramas, com todas as formas e representações e possibilidades de um termo.

Quanto a ações para desenvolvimento dessas ontologias robustas, foi exposto que para o campo de produção científica, nos formatos mais tradicionais (artigos, teses e dissertações), ainda não há iniciativas no desenvolvimento dessas ontologias pois o GT de indexação está debruçado em outras ações, como a elaboração da Política. No entanto, para dados de pesquisa isso já está em andamento. No âmbito da Rede de Bibliotecas há um grupo que participa de forma articulada com o *Go-Fair* Brasil Saúde, que tem um GT voltado para ontologias. Atualmente está sendo desenvolvida uma ontologia para Zika, com participação de um pesquisador renomado da Fiocruz que está apoiando e validando esses descritores. Eles estão trabalhando numa perspectiva de ontologia de fundamentação por áreas de pesquisa para facilitar a indexação nos diretórios tanto por bibliotecários de dados, quanto do próprio pesquisador. À princípio foram pensadas como piloto as áreas da Zika e pneumologia, mais especificamente tuberculose, mas em virtude da pandemia, houve a necessidade de priorizar e, por isso, foi escolhida a Zika. Cabe ressaltar

que há uma forte possibilidade de entrar com a COVID-19 também, porque há um grande desenvolvimento internacional neste sentido.

## II. Biblioteca da Saúde da Mulher e da Criança

A Biblioteca da Saúde da Mulher e da Criança (BibSMC) está localizada no Instituto Fernandes Figueira (IFF), hospital da Fiocruz, no bairro do Flamengo. Atende aos pesquisadores, professores, alunos, profissionais do IFF (farmacêuticos, enfermeiros, médicos, nutricionistas...) e público externo.

Tem a missão de desenvolver novos métodos, processos e produtos para ampliar e universalizar o acesso à informação científica na área da saúde da mulher, do adolescente e da criança. [...] Áreas de Especialização: Aleitamento Materno, Genética, Ginecologia, Infecção Hospitalar, Nutrição, Obstetrícia, Patologia, Saúde Mental. Especialidades Pediátricas: Alergia, Cardiologia, Cirurgia, Clínica, Doenças Infecciosas, Gastroenterologia, Hematologia, Neonatologia, Neurologia, Patologia, Pneumologia, Radiologia, Tratamento Intensivo (LEONARDO, 2008).

A Biblioteca conta com quatro funcionários. Usam a Classificação Decimal de Dewey (CDD) e os Descritores em Ciências da Saúde (DeCS) como Sistemas de Organização do Conhecimento (SOC). A CDD está em formato físico e o DeCS *on-line*.

Segundo o profissional entrevistado, o sistema de gerenciamento de bibliotecas Aleph foi adotado para integração das bibliotecas da Rede, mas foi respeitada as particularidades de cada biblioteca. De modo geral, todos se voltam para um diálogo com o DeCS. O DeCS é um vocabulário controlado para uniformização dos termos na área de saúde criado pela Bireme e baseado no Medical Subject Headings (MeSH), que é a linguagem usada pela Medline/PubMed, da Biblioteca Nacional de Medicina dos EUA. O profissional lembra que, com a intenção de fazer algo similar na América Latina, a Bireme e OPAS desenvolveram a base de dados LILACS, para reunir a produção Latino Americana em Ciências da Saúde que utiliza os descritores do DeCS. No DeCS pode-se fazer buscas em inglês e português e há uma uniformização dos termos na área de saúde. Para escolha dos seus descritores, é necessário entrar no seu *site* de maneira externa, pois não está implementado no Aleph, que é o sistema de gerenciamento usado pela biblioteca.

O profissional considera importante a demarcação de metadados e a interoperabilidade entre sistemas para facilitar a recuperação da informação. Por isso o DeCS facilita, pois dialoga com outras bibliotecas da área da saúde. Então, para

alimentar o Aleph e o repositório institucional Arca, com a produção formal do IFF, os descritores devem ser baseados no DeCS. Acredita que isso trará maior consistência na recuperação dos dados.

Não considera existir dificuldades em seu uso, pois os descritores atendem bem à temática abarcada pela biblioteca. Lembra ainda que o DeCS é atualizado continuamente, apesar de não ser uma atualização muito rápida, mas tem um corpo de profissionais dedicados para revisão e inclusão de novos termos, inclusive já há uma equipe estudando sobre o Covid-19. Há todo um trâmite até que o termo seja consolidado e incorporado ao DeCS. Considera ser uma ferramenta dinâmica, dentro dos seus limites.

Com relação à recuperação da informação, não identifica problemas, e informou que a biblioteca oferece serviço de referência, com bibliotecários que assessoram os usuários na pesquisa, para construir juntos uma estratégia de busca. Assim expandem e mapeiam os materiais no acervo da biblioteca e também em bases da Fiocruz, portal Capes e outras. É um trabalho mais elaborado, construído junto com o usuário.

Quanto aos projetos em desenvolvimento, afirmou haver discussão na Fiocruz sobre ciência aberta e mencionou o Projeto Educare, um projeto bem novo que nasceu pouco antes da pandemia, relacionado ao campus virtual da Fiocruz. O profissional acredita que podem desenvolver algum projeto de integração, com acesso aberto ao conhecimento, de forma que funcione como um filtro e ajude na recuperação. A Biblioteca também promove treinamentos para uso de bases de dados para os profissionais do IFF, fazendo um trabalho de educação de usuários.

### **III. Biblioteca de Educação e Divulgação Científica Ilone Seibel**

A Biblioteca de Educação e Divulgação Científica Iloni Seibel tem seus serviços voltados aos alunos dos cursos oferecidos pelo Museu da Vida. Suas atividades iniciaram em 1999 prestando apoio informacional aos profissionais do Museu da Vida, no desenvolvimento de projetos educacionais, atividades e eventos voltados ao público visitante (MUSEU DA VIDA, 2020?).

Além de atender aos profissionais da Fiocruz, a sala de consulta é aberta ao público externo, com acesso local ao seu acervo e ao Portal de Periódicos da Capes. Dispõe de sala de vídeo, que mediante agendamento por qualquer pessoa, viabiliza a exibição de sua coleção de filmes, animações e documentários educativos sobre ciências, meio ambiente e saúde.

O acervo possui, aproximadamente, 5 mil itens, nacionais e estrangeiros, nas áreas de divulgação científica, educação, museologia, ciências

da vida, meio ambiente, saúde e literatura infantojuvenil, sendo formado por livros, folhetos, teses, dissertações, periódicos, CD-ROM, DVD, jogos e obras de referência. Sua coleção infantojuvenil tem o objetivo de divulgar a ciência junto ao público visitante do Museu, por meio de atividades de incentivo à leitura, articuladas a temas de ciência e saúde (MUSEU DA VIDA, 2020?).

Segundo o profissional entrevistado, diante da especificidade da Biblioteca, existia um desafio para realizar a indexação dos assuntos do seu acervo bibliográfico, pois além da CDD e do DeCS (Descritores em Ciências da Saúde) era necessário pesquisar em outras fontes em função da limitação dos termos e consequentemente, para dar conta de temas como: Museologia, Literatura Infantojuvenil, divulgação científica, ou seja, um acervo multidisciplinar, optou-se em consultar outros catálogos que pudessem apoiar uma classificação mais fidedigna. Os catálogos utilizados como fonte foram o da Biblioteca Nacional (BN), da Library of Congress, e o de bibliotecas universitárias como USP e UNICAMP.

O profissional ressalta que durante essa pesquisa em outros catálogos, iniciou-se na Unidade, um Grupo de Trabalho (GT) em Indexação preocupado com o vocabulário controlado que atendesse o acervo da Casa de Oswaldo Cruz (duas bibliotecas distintas, arquivo, reserva técnica, museologia e o patrimônio histórico e arquitetônico). Neste GT ficou definido que seria feito um levantamento de todo o acervo com o objetivo de criar um vocabulário controlado que atendesse a todo acervo da Casa de Oswaldo Cruz. Foi realizado após este levantamento, um “saneamento” e/ou “Higienização” dos termos, de maneira que diminuísse possíveis erros na busca e a recuperação da informação tenha mais qualidade. Posteriormente foi feito o agrupamento dos assuntos de todos os acervos e criado o vocabulário controlado.

Outro ponto de destaque reforçado pelo profissional entrevistado, é que as Bibliotecas da Fundação Oswaldo Cruz utilizavam diferentes ferramentas de catalogação e classificação do acervo bibliográfico. Após 5 anos, por meio da Rede de Bibliotecas foi decidido a utilização do Aleph como plataforma para todas as bibliotecas da rede, nomeada na Fiocruz como Catálogo Mourisco, facilitando assim a recuperação da informação pelo usuário.

**Quadro 4 - Lista de Áreas Temáticas da Coleção Infantojuvenil.**

ADOLESCÊNCIA	CORES	MITOS
ARTES	DIVERSIDADE	MOSQUITO
CÉREBRO	EXPERIMENTOS	MÚSICA
CIÊNCIA	IMAGEM	POESIA

**Quadro 4 - Lista de Áreas Temáticas da Coleção Infantojuvenil.**

CLÁSSICOS	LIVRO	POVOS INDÍGENAS
COMPORTAMENTO	MEIO AMBIENTE	SAÚDE E HIGIENE
COOPERAÇÃO	MEMÓRIA	TEATRO

Fonte: Biblioteca de Educação e Divulgação Científica Iloni Seibel (2020).

Torna-se importante enfatizar que a Biblioteca de Educação e Divulgação Científica, como relatou o profissional, tem especificidades por atender um público infantojuvenil e temas como: diversidade, experimentos, saúde e higiene, mosquito, entre outros para atender a demanda dos contadores de história. Sendo assim, além da CDD e Tabela de Cutter, eles sentiram a necessidade de utilizar outra ferramenta para a classificação, e foi elaborado uma lista temática para reunir a coleção infantojuvenil (Quadro 4) e organizar por cores, para facilitar o manuseio das crianças e adolescentes. A lista também é útil ao propósito dos contadores de história ao planejar as atividades internas da biblioteca ou do Museu da Vida.

Outra forma de organização do acervo bibliográfico desta Biblioteca é a classificação das dissertações e monografias de conclusão de curso, dos quais a biblioteca é responsável pela guarda. A estratégia utilizada foi ordenar pelas primeiras letras dos nomes do autor e ano da conclusão do curso.

#### **IV. Biblioteca de Histórias das Ciências e da Saúde**

O profissional entrevistado relatou que, por atuar em rede, existe um padrão mínimo que seguem, por isso utilizam para catalogação a Classificação Decimal de Dewey (CDD) que atende plenamente as necessidades desta biblioteca e a Tabela de Cutter também é utilizada na complementação de número de chamada.

Segundo o profissional, para indexar, considerando que o DeCS não contempla todas as áreas temáticas em História da Saúde, foram desenvolvidos descritores locais para dar conta das especificidades da indexação. É importante sinalizar que para a construção desses descritores locais foram consultadas outras terminologias como a lista de assuntos da Library of Congress, da Biblioteca Nacional (BN), o Tesaurus Spines e a descontinuada base Thermus, que armazenava no LILDBI os descritores locais instituídos no correr dos anos.

Na migração para o Sistema do Aleph, com perspectiva de rede e a gestão bibliográfica como a catalogação e indexação, está sendo realizada no novo sistema a criação de um módulo para alimentar os descritores para fazer o controle das autoridades. Em relação ao uso da ferramenta, a Biblioteca ainda está em adaptação devido à pandemia.

No ano de 2016, para contemplar todos os acervos e criar a interoperabilidade dos descritores entre o museu, biblioteca, arquivo e acervo arquitetônico foi rea-

lizado um mapeamento de todos os descritores das tipologias dos acervos e posteriormente foi criada uma lista, atualmente em *excel*. Apesar de não estruturada, essa lista tem sinônimos dos descritores de assuntos e também possui descritores de autoridades que será chamado de vocabulário controlado, pois tem o objetivo de unificação.

A Unidade possui mais de 60 coleções, algumas de doação, que são classificadas pela CDD, porém são agrupadas fisicamente separadamente.

O público-alvo são os alunos de Mestrado e Doutorado em História de Ciências da Saúde e alunos do Mestrado em preservação do patrimônio, professores, pesquisadores e público externo.

#### **V. Biblioteca de Manguinhos**

A Biblioteca de Manguinhos integra a Rede de Bibliotecas da Fiocruz, e atende principalmente ao Instituto Oswaldo Cruz (IOC), mas também a quatro outras unidades, a saber: o Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICTCT); Instituto Nacional de Infectologia Evandro Chagas (INI); o Instituto de Ciência e Tecnologia em Biomodelos (ICTB); e o Instituto de Tecnologia em Imunobiológicos (Bio-Manguinhos).

Tem por missão desenvolver novos métodos, processos e produtos para ampliar e universalizar o acesso à informação científica na área biomédica. Suas ações são destinadas, especialmente, aos profissionais de saúde, alunos de pós-graduação, professores e pesquisadores da Fiocruz, das redes pública e privada de saúde, atendendo também a sociedade em geral (MAGALHÃES, 2017).

É especializada na área de biomédica, cobrindo as áreas de “Bacteriologia, Biologia, Biologia Molecular, Bioquímica, Biotecnologia, Entomologia, Farmacologia, Genética, História Natural, Imunologia, Medicina Tropical, Medicina Veterinária, Micologia, Microbiologia, Microscopia, Parasitologia, Patologia, Virologia, [e] Zoologia” (MAGALHÃES, 2017).

Foram declarados os usos dos seguintes SOC: o DeCS como recurso externo, sendo este utilizado para toda a Rede de Bibliotecas da Fiocruz; e a Classificação Decimal de Dewey (CDD), utilizada na grande maioria das Unidades de Informação da rede.

Também usa-se localização fixa para coleções especiais que foram institucionalizadas. Fruto de doações de pesquisadores ou instituições que doaram seus Centros de Documentação. A localização fixa foi selecionada com a ideia de pre-

servar a identidade das coleções recebidas, mantendo assim a classificação atribuída aos itens em sua origem, ou seja, a classificação feita pelo doador enquanto proprietário.

Declarou-se que a instituição tem um tesouro próprio, mas não está em uso em função de desatualização. O tesouro foi desenvolvido inicialmente para o acervo da Biblioteca de Manguinhos, e dada esta íntima relação, resolveu-se explorá-lo.

Conforme relato do profissional entrevistado, o Tesouro de Manguinhos (Teman) foi criado em 1986, a partir de projetos financiados pela FINEP, tendo como objetivo subsidiar o processo de indexação da Biblioteca de Manguinhos. Sua estrutura foi pensada inicialmente para atender a demanda especializada do Instituto Oswaldo Cruz, e teve como base o periódico Memórias do Instituto Oswaldo Cruz. O periódico foi a matéria prima essencial para que, a partir de seu conteúdo, fossem definidas as famílias, os gêneros, as ordens e toda a terminologia que compõem o tesouro, ou seja, sua estrutura.

O trabalho foi realizado por um grupo de pesquisa que foi instituído para identificar e analisar todo o conteúdo de todos os exemplares desta revista periódica. Posterior ao trabalho inicial de identificação e análise, realizou-se uma filtragem para a composição dos termos do Teman.

O Teman foi muito utilizado para indexar o projeto Produção de Manguinhos (PROMAN). O PROMAN consiste na reunião de toda produção científica dos pesquisadores da instituição, são arquivos reproduzindo a produção científica de todos eles, desde Oswaldo Cruz e os demais pesquisadores fundadores da instituição.

O acesso ou consulta ao Teman se dá de forma digital no próprio sistema de gerenciamento de biblioteca, o sistema Aleph. Não existe uma publicação física ou arquivo digital (.pdf por exemplo) do mesmo. A visualização só é possível pelo indexador, ao *logar* o sistema com seu perfil de operador da plataforma.

Todavia, mesmo durante o processo de indexação pelo uso do Teman, o indexador não tem acesso à sua estruturação hierárquica, embora ela esteja estruturada na linguagem do sistema. Sua apresentação na interface é uma lista em ordem alfabética.

No uso do módulo de catalogação do Aleph, ao dar início ao processo de indexação, o bibliotecário pode fazer a busca por palavras e selecionar a base de dados Tesouro de Manguinhos (FIO01), para pesquisar entre os termos desta estrutura. Uma lista que somatiza os resultados obtidos é ofertada para seleção do bibliotecário. Após acessar a lista “resultado”, apresentam-se listados todos os termos que foram recuperados pela busca.

O Teman é tratado dentro do Aleph como uma base de dados; uma base de dados de indexação. E por este termo “base de dados” são tratadas quaisquer lin-



guagens de indexação que venham a ser implementadas no sistema, fazendo-se necessário que o indexador selecione na interface de operação, qual base vai utilizar naquele momento para buscar um termo. É possível indexar um mesmo item a partir de diversas bases, bastando para isso trocar a base e realizar nova busca de termos existentes. O registro final do item teria termos indexados de diferentes bases, promovendo diversidade na forma de recuperação para cada item.

Além da situação de desatualização, apontou-se como característica que oferece dificuldade na indexação com este tesouro, o fato do Teman não apresentar via sistema, uma hierarquização visual para o indexador perceber, durante o processo de indexação, o nível de especificidade que o termo recuperado para a indexação representa dentro da estrutura. Foi considerado como uma característica essencial ausente.

Uma outra característica relatada como dificuldade é a existência de evidências de linguagem natural na estrutura do Teman. Quanto à percepção sobre a recuperação da informação durante a pesquisa, foi demonstrado o processo no sistema Aleph, evidenciando que o software ainda precisa de customizações para oferecer maior riqueza de possibilidades ao pesquisador. A implantação ainda está em processo, e as otimizações são objetivos traçados para oferecer robustez à ferramenta quanto ao seu poder de recuperação.

Na interface do usuário pesquisador, por exemplo, ainda não é possível utilizar como filtro para campo de assunto uma lista específica que esteja implementada no Aleph, seja ela um tesouro, um cabeçalho de assuntos, ou qualquer outra Base de Dados de indexação. Mas, estando a indexação realizada de forma a identificar a base que serviu de fonte, o pesquisador poderá ter o conhecimento de sobre qual base aquele termo foi indexado no item recuperado.

A estrutura do sistema OPAC em vigor é muito antiga, não permitindo uma recuperação para os usuários com resultados tão ricos quanto poderia ser ofertada com a nível de catalogação e indexação realizado no sistema. Ainda não se pode criar equações de busca com operadores *booleanos* que dêem conta de todos os aspectos e metadados gerados pela catalogação ou indexação que a unidade realiza, o que torna limitado, os recursos de autonomia dos usuários. Um trabalho atual está sendo realizado para otimizar o sistema OPAC para oferecer mais riqueza de opções à busca do usuário.

Além da remodelação do OPAC para oferecer mais riqueza de opções à busca do usuário, foi apresentado como projeto futuro, a intenção de aprimorar o Teman via esforço do grupo de pesquisa. Para atualizá-lo é preciso reunir um grupo de pesquisa multidisciplinar, com diversos representantes do IOC, nas diversas áreas do conhecimento que a unidade abrange.

Para as outras unidades que possuem um outro eixo temático, o Teman seria útil apenas no esquema da área biomédica.

#### **VI. Biblioteca de Medicamentos e Fitomedicamentos**

A Biblioteca de Farmanguinhos, como é conhecida, tem como áreas de especialização: Pesquisa e Desenvolvimento de Produtos Farmacêuticos; Indústria farmacêutica; Avaliação de produtos farmacêuticos; Prospecção de produtos na indústria farmacêutica; Inovação em Fitomedicamentos; Gestão Tecnológica da Pesquisa e Desenvolvimento na indústria farmacêutica e Pesquisa e Desenvolvimento para Prospecção de Produtos Farmacêuticos. Integra o Complexo Tecnológico de Farmanguinhos, estando vinculada ao Instituto de Tecnologia em Fármacos Farmanguinhos. Conta com três funcionários na biblioteca, sendo somente uma pessoa envolvida com os soc. Atende alunos de especialização, profissionais da unidade, outras bibliotecas e também é aberta para consulta ao público externo.

O profissional entrevistado informou que, com a intenção de integrar todas as áreas de informação da unidade, criaram o Departamento de Gestão da Informação e Documentação que inclui a biblioteca e o arquivo, para uma melhor gestão de documentos.

Seguem a linha da rede Fiocruz e utilizam a CDD (*on-line*) e o DeCS (*on-line*), possuem ainda uma planilha de termos que remetem ao DeCS, chamado de vocabulário controlado, de uso exclusivo desta biblioteca. Usam também a tabela Cutter de notação de autores, para compor o número de classificação.

O profissional considera o DeCS muito bom para área de saúde, mas ele não abarca alguns termos relacionados à indústria farmacêutica. Apontou algumas dificuldades na busca dos termos, como palavras conhecidas e usadas no cotidiano que muitas vezes são difíceis de serem recuperadas. Deu como exemplo o termo “Medicamentos”, que no DeCS é uma remissiva para “Preparações farmacêuticas”. Além de não ser usado o termo mais conhecido como autorizado, a busca por remissivas também não é muito precisa.

Com relação à CDD, alguns assuntos específicos não são atendidos muito bem, mas de forma geral não há maiores problemas. A CDD oferecia dificuldade enquanto impressa por ser uma versão desatualizada. A aquisição da versão *on-line* suprimiu a deficiência de desatualização, mas a renovação estava pendente até o momento da entrevista. O profissional inteirou que a rede de bibliotecas pretende fazer assinatura para todas as unidades terem acesso.

Informa que as dificuldades para indexação e classificação estavam mais nas ferramentas de trabalho: CDD antiga e desatualizada, e o DeCS com termos que

não atendiam completamente. Assim surgiu a necessidade da criação de um vocabulário controlado próprio, com os termos mais usados. Foi desenvolvido há mais de 10 anos por bibliotecários pregressos, especificamente para o contexto de Farmanguinhos.

Esse vocabulário funciona como uma lista de remissivas para auxiliar a busca no DeCS. Consiste em uma planilha com os termos mais conhecidos remetendo aos termos autorizados pelo DeCS. Está em formato impresso e só é usada pela biblioteca de Farmanguinhos, não é algo compartilhado com outras unidades.

O profissional entrevistado lembra ainda que algumas informações são documentadas em um livro para registro de decisões e informações para futuros funcionários. Possuem também uma planilha de controle quantitativo do acervo. E alimentam o repositório institucional Arca, com produções da unidade e a Revista Fitos.

Foi comentado também sobre a recente adesão ao Aleph (sistema de gerenciamento de bibliotecas), que ainda está em fase de correção e adaptação, pois a migração ocorreu pouco antes da pandemia. Sobre a recuperação em seu acervo ainda não há feedbacks, mas com o sistema antigo, o LILDBI, não existiam dificuldades.

## **VII. Biblioteca de Saúde Pública**

A Biblioteca de Saúde Pública localizada na Escola Nacional de Saúde Pública Sérgio Arouca (ENSP) conta com 10 funcionários, dos quais 3 estão envolvidos no uso dos Sistemas de Organização do Conhecimento. Seu acervo abrange as áreas de saúde pública e saúde mental e atende aos pesquisadores alunos e funcionários da ENSP, comunidade acadêmica em geral e usuários externos.

Como na maioria das bibliotecas da Rede, utiliza a CDD 23ª edição (impressa) para classificar e o DeCS (*on-line*) para indexar. O profissional entrevistado reitera que há semelhanças com outras bibliotecas da Fiocruz, mas cada uma tem seu jeito de trabalhar, suas particularidades.

O profissional informou que usavam o sistema LILDBI para gerenciamento do acervo, mas recentemente houve a migração para o sistema Aleph. O objetivo é que as bibliotecas da Rede Fiocruz estejam integradas. Com isso, o usuário poderá fazer uma única carteirinha e usar em todas as bibliotecas da rede.

Ainda estão em período de adaptação, uma vez que o LILDBI era mais simples. Já o Aleph é mais detalhado e usa o formato Marc, o que facilita algumas tarefas, como por exemplo, importar os dados de catalogação de outra biblioteca, agilizando o processo. Mesmo quando é realizada essa importação de dados, fazem uma análise dos assuntos para atender aos seus usuários e especificidades. Enquanto as teses e dissertações são muito específicas de cada unidade, então a catalogação é manual.

O profissional ressalta ainda que quanto mais especializada é sua biblioteca, mais se torna necessário especificar os assuntos. Pois uma boa indexação reflete diretamente na recuperação e ajuda principalmente para quem não é da área.

A dificuldade a considerar no DeCS refere-se na demora para atualizar os descritores, face a existência de uma comissão para validação. A CDD também oferece lentidão de atualização e ocorre casos de não existir notação mais específica, o que força a classificação em número mais geral. Pretendem adquirir acesso a *Web Dewey* (versão da CDD *on-line*) para acompanhar a atualização de forma mais rápida. De toda forma, essas dificuldades não atrapalham o serviço da biblioteca.

A forma de armazenamento adotada é pelo número da CDD. Os usuários têm acesso livre às estantes no 1º andar da biblioteca onde fica o acervo geral, mas no 2º andar, onde ficam as teses e dissertações, precisam da ajuda dos funcionários. Alguns funcionários, que não são da área, têm dificuldade para localizar pelo número da CDD, então existe uma pretensão em mudar as teses e dissertações para outro tipo de localização, como a fixa, de forma que sejam encontradas com mais facilidade, mesmo por quem não é bibliotecário.

Com relação à recuperação dos itens, o profissional entrevistado acredita que os sistemas de gerenciamento podem influenciar na busca. O LILDBI tinha possibilidade de indexar com as palavras-chave dos autores, mas no Aleph, ainda não tem campo para isso. Como a migração ocorreu recentemente e veio a pandemia logo depois, não tem muitas informações sobre essa recuperação. O sistema está sendo alimentado com os metadados, mas ainda não está disponível para busca e recuperação. Pretendem fazer inventário para somente depois disponibilizar aos usuários.

Com a mudança de sistemas, perceberam que havia muitos itens semelhantes, mas com assuntos diferentes, pois a indexação é um pouco subjetiva, depende do olhar do indexador. Por isso, pretendem desenvolver um documento com diretrizes para uniformização da indexação e diminuir a subjetividade do processo, pois é importante seguir um padrão na instituição. Estão revendo os processos de catalogação e indexação e querem formalizar essas instruções em documentos, como uma política de indexação, por exemplo. Enquanto a biblioteca está fechada, querem sanar esses problemas, para oferecer um trabalho melhor para os usuários.

Para fazer a ficha catalográfica de teses e dissertações é necessário explicar para os autores que nem sempre existe o termo sugerido no DeCS. Alguns autores querem que sejam usadas as suas palavras-chave e em determinadas ocasiões, é difícil fazê-los entender a necessidade de adaptar para os descritores controlados. Então, quando não encontram o termo exato no DeCS, usam um aproximado. O profissional lembra que é importante ver uma forma de atender a essa demanda dos usuários, pois o serviço da biblioteca é voltado para eles. Existe também a questão

de que alguns termos usados pelos acadêmicos, podem ainda não ter sido inseridos no DeCS. Acredita que é possível ser mais flexível e chegar a um equilíbrio.

O profissional entrevistado mencionou sobre a intenção de que as fichas catalográficas possam ser feitas *on-line*, de forma que as palavras-chave sejam integradas com os descritores do DeCS. Isso ajudaria nos problemas dos autores que desejam usar as próprias palavras-chave, pois dessa forma eles já fariam a busca baseada nos termos autorizados.

#### **VIII. Biblioteca Emília Bustamante (BEB)**

A Biblioteca Emília Bustamante fica na Escola Politécnica de Saúde Joaquim Venâncio (EPSJV), a qual oferece cursos técnicos de nível médio, de especialização e de qualificação nas áreas de Vigilância, Atenção, Informações e Registros, Gestão, Técnicas Laboratoriais, Manutenção de Equipamentos e Radiologia, além da Educação de Jovens e Adultos (EJA) e de um Programa de Pós-graduação em Educação Profissional em Saúde. A biblioteca conta com 8 pessoas na equipe e tem um perfil mais voltado para os alunos do nível médio, no entanto também atende alunos da pós-graduação.

Utilizam a CDD 22<sup>a</sup> ed. (impressa) para classificação dos materiais bibliográficos. A indexação é feita usando os descritores do DeCS para área de saúde. Já para a área de Educação, usam o tesauro do INEP. A prioridade é usar o DeCS, mas se necessário, recorrem a outras bases. A indexação na BVS exige o uso do DeCS, pois é o padrão da Bireme.

Segundo o profissional entrevistado, não sentem dificuldades no uso da CDD. Pretendem comprar a 23<sup>a</sup> ed, mas há intenção de assinar a *Web Dewey* (versão da CDD *on-line*) para toda rede. O DeCS tem um processo para atualização, o que implica na demora para inclusão de termos novos, por isso precisam fazer pequenas adaptações, usando termos aproximados. Por exemplo, na elaboração de fichas catalográficas da pós-graduação, é necessário essa adaptação para termos solicitados pelos autores, que porventura ainda não existam no DeCS.

O sistema de gerenciamento de bibliotecas ainda é o PHL. A versão paga do PHL é atualizada com frequência e há bom contato direto com o desenvolvedor. Na EPSJV existem algumas particularidades que o PHL atende muito bem, não há reclamações. Uma vantagem do PHL é que podem usar as palavras-chave dos autores e também os termos do DeCS. Todo acervo está no PHL, exceto a BVS, que é bem particular.

Houve a migração do acervo para o Aleph, que está em fase de revisão, mas ainda não está sendo usado. O profissional reconhece que a unificação da base para todas as bibliotecas facilita a comunicação e reduz outros problemas. Essa mudança para o Aleph está sendo feita aos poucos para melhor adaptação.

Ainda não usam o Aleph para busca e recuperação, mas o PHL atende bem. O profissional lembra que o PHL tem a possibilidade de incluir a imagem da capa do livro, o que facilita na busca, pois muitos alunos procuram o livro pela cor da capa. A maior parte dos usuários são alunos do nível médio escolar, e não têm o hábito de pesquisar por nomes de autores, por exemplo.

O profissional entrevistado comentou sobre o Grupo de Trabalho de Normalização da Fiocruz que discute a padronização para toda rede, para usar os mesmos instrumentos e ferramentas de trabalho. A implantação do Aleph já é uma tentativa.

Na biblioteca Emília Bustamante está sendo criado um pequeno acervo infantil, mas ainda em fase inicial, para atender a demanda de um grupo de crianças que são os filhos de estudantes da modalidade da Educação para Jovens e Adultos (EJA). É usado o espaço da biblioteca para receber as crianças enquanto os pais estudam e há dois profissionais moderadores que fazem o acompanhamento. Pretendem fazer um projeto para atender a essas crianças e precisam pensar na forma (ou SOC) a ser utilizada para organização desse acervo, ainda em fase inicial de desenvolvimento.

#### **IX. Biblioteca Virtual Oswaldo Cruz**

De acordo com o profissional entrevistado, a criação da Biblioteca Virtual Oswaldo Cruz se deu a partir da necessidade de otimizar a disponibilização dos documentos existentes nos diversos acervos da COC. Por ser a única unidade da Fiocruz que possui outros acervos, além do bibliográfico, a COC teve pioneirismo nessa unificação de acervos, por processo colaborativo, através de técnicas da Ciência da Informação, por Sistemas de Organização do Conhecimento.

Os acervos arquitetônico, arquivístico, bibliográfico e museológico dessa unidade, possuem diferentes tipologias documentais que, anteriormente, não conversavam entre si, embora tivessem a mesma linha temática: História das Ciências e da Saúde. Isso despertou a necessidade de uma ação que permitisse conexão entre esses diferentes acervos já que possuem pontos em comum como estarem dentro de uma mesma unidade técnico-científica, possuírem o mesmo conjunto de conceitos e estarem organizados por profissionais documentalistas preocupados com a entrada e saída de dados. Cabe destacar que cada acervo está organizado por uma classificação própria, ou seja, o acervo bibliográfico tem um sistema de classificação diferente do utilizado no acervo arquivístico, que por sua vez é diferente do museológico, do arquitetônico, fora os acervos que já estavam em formato digital que também tinham um sistema de organização específico.

Assim, a forma encontrada para essa interação entre os acervos e suas respectivas formas de classificação e bases de dados foi a indexação. A partir desse entendimento, foi iniciado o Projeto de integração, tendo como piloto a Biblioteca Virtual de Oswaldo Cruz, utilizando a indexação não apenas como base para registros de seus metadados como também da própria gestão de conteúdo deste *site*.

A Biblioteca Virtual de Oswaldo Cruz não existe fisicamente. Apresenta-se, exclusivamente, em formato *on-line* e pode ser acessada pelo *site* no endereço: <http://oswaldocruz.fiocruz.br/index.php>. Trata-se de uma recriação de todos os acervos existentes que foram unificados e, portanto, a sua organização é apenas virtual. A biblioteca

Oferece um panorama da vida do cientista, com ênfase em sua trajetória profissional, baseada em uma variada documentação histórica, oriunda do arquivo pessoal, da coleção bibliográfica, e de peças museológicas que se encontram sob a guarda Casa de Oswaldo Cruz (FUNDAÇÃO OSWALDO CRUZ, [201-?]).

Sua criação tem por finalidade que o usuário tenha acesso, em uma única busca, à reunião de toda informação disponível sobre o assunto pesquisado, independentemente do acervo ao qual pertença. Exemplo: o usuário tem o objetivo único de pesquisar uma planta baixa de um prédio da Fiocruz. Com essa busca integrada, ele pode ter acesso aos documentos administrativos relativos à sua construção e manutenção, às informações orçamentárias, às fotografias do prédio e eventos nele realizados e muito mais, propiciando, segundo o profissional entrevistado, a ampliação de seu escopo de pesquisa.

Conforme destaca o profissional, essa Biblioteca Virtual é uma biblioteca de referência, não fazendo parte das BVS da Bireme, porque não houve outro desenvolvedor parceiro, senão a COC, para colaborar e agregar. Ela foi nomeada como Biblioteca Virtual mas, na verdade, é um *site* bibliográfico específico para a Casa de Oswaldo Cruz. Funciona como uma interface de busca que pode redirecionar à base que contém o assunto pesquisado ou, se o pesquisador não quiser ir ao banco de dados de origem para obter maiores detalhes, pode fazer a visualização pretendida no próprio *site*. A equipe envolvida na sua manutenção e aprimoramento abrange profissionais de todos os acervos da COC (arquitetônico, arquivístico, bibliográfico, museológico e virtual).

O profissional relata que a principal forma de organização do conhecimento dessa biblioteca é a indexação. Para essa base foi criada uma nova indexação, independente, ou seja, se o resultado da busca aponta para algum documento que está

na base de dados do Arquivo Nacional (AN), respeita-se o descritor que está na base original do AN, mas na Biblioteca Virtual há uma nova lista de assunto, com uso do DeCS e dos descritores locais.

Há muitos descritores locais, palavras-chave, termos, que dizem respeito aos acervos da COC e o uso do DeCS não atende plenamente porque a linha temática dessa biblioteca não é, exclusivamente, Ciências da Saúde. Por isso, para unificar e integrar os diferentes acervos, foi preciso fazer uma nova indexação alinhando duas listas de termos: uma já existente, o DeCS, e uma nova lista que foi definida com os termos locais utilizados nesses acervos.

Como o profissional entrevistado é membro do Grupo de Trabalho de Indexação da Rede de Bibliotecas da Fiocruz, contextualizou que sua participação representando a COC se dá pelo fato da criação da Política de Indexação de Acervos e o Manual de Indexação daquela unidade, desenvolvido de maneira que mesmo profissionais não bibliotecários consigam utilizar adequadamente, o que demonstra a facilidade de seu uso.

Além disso, com a Política e o Manual, a COC criou uma metodologia chamada saneamento, onde foram analisadas as listagens de termos já indexadas nas bases de dados dos acervos, o que gerou uma lista alfabética de assuntos com os termos mais empregados em todos os acervos da COC. Essa lista é chamada de Vocabulário Controlado da COC, pois a intenção é que se torne um, mas ainda se configura como um cabeçalho de assuntos.

Há a perspectiva, vislumbrada desde o escopo de produtos do projeto inicial, de que se tenha uma ferramenta própria que transforme essa lista em algo semelhante ao DeCS. Atualmente, ela é uma planilha, disponível *on-line* no *Google Drive* para permitir compartilhamento e consulta. Como destaque do próprio profissional entrevistado, essa lista foi produzida de forma participativa e democrática entre todas as áreas envolvidas, visto a intenção de que os profissionais dos acervos pudessem se reconhecer nos termos apresentados e que isso lhes propiciasse independência no ato de indexar.

O profissional ressalta, ainda, que a iniciativa da COC em desenvolver seu próprio vocabulário controlado, mesmo com a existência do DeCS, se deu com o propósito de abarcar os termos do DeCS mas também o que não está lá, porém é utilizado pelos demais acervos da COC. Esse cabeçalho de assuntos fortalece o uso de padrão para indexar os termos, diminuindo a subjetividade no processo de indexação.

Conforme expressado pelo profissional, a Rede de Bibliotecas está seguindo o caminho trilhado pela COC a partir da demonstração do bom funcionamento da estrutura de indexação como Sistema de Organização do Conhecimento. Isso porque o Aleph tem uma lista de assuntos acoplada e permite que novos termos sejam subscri-



tos pelo usuário quando ele descreve e o termo ainda não existe na base. No entanto, isso carrega subjetividade e possibilita erros que acabam por gerar novos termos, “sujando” essa lista o tempo todo. Desta forma, ressalta a importância de uma Política e um Manual, que funcione de maneira bem intuitiva e que demonstre a necessidade de que os qualificadores sejam usados de forma consciente. Ou seja, o SOC visa facilitar a vida do usuário e bibliotecário no momento de indexação e recuperação.

#### **X. Repositório Institucional ARCA**

O Arca é um sistema, um repositório, sob plataforma DSpace de código aberto, criado com a intenção de reunir em um único lugar toda produção da Fiocruz. Toda produção realizada por servidor, bolsista, estagiário, e até terceirizado, enquanto vinculado à Fiocruz, entra para o Arca.

Segundo o profissional entrevistado, o DSpace é usado por 80% das instituições nacionais e internacionais para criar repositórios institucionais. Têm-se definidos alguns metadados em Dublin Core que “conversam” por interoperabilidade com os metadados estruturados em MARC no sistema Aleph das bibliotecas.

A indexação realizada no Arca tem por base o DeCS por ser o vocabulário controlado e autorizado para descritores da área de saúde. O DeCS é baseado no Medical Subject Headings (MeSH).

Todavia, em função da automação de importação dos dados inseridos pelos bibliotecários no sistema Aleph, os descritores indexados com base em quaisquer outros tesouros ou vocabulários que não seja o DeCS também são importados e mantidos dentro do Arca como indexação do item importado do Aleph. Neste sentido, a indexação realizada no Aleph pelos bibliotecários em cada uma das bibliotecas é respeitada.

O profissional lembra que no Arca, os termos de indexação são identificados como palavras-chaves, e podem ser com base em qualquer linguagem controlada ou até mesmo em linguagem natural. Depende apenas da prática que tenha sido realizada em cada Unidade de Informação. Informa que nos casos de teses, dissertações e artigos, são consideradas as palavras-chaves elaboradas pelos bibliotecários da unidade que recebeu o depósito e considerou se as palavras chaves definidas pelos próprios autores são pertinentes.

Em resumo, além dos metadados básicos como título, autor, data, resumo, referências, notas avulsas etc, há campos para receber metadados de indexação do DeCS e palavras-chaves distintamente.

Um outro recurso implantado no Arca para oferecer gerenciamento destas informações, é a ferramenta “visualização de dados”, destinada a minerar as informações permitindo que se possa, por exemplo, conhecer qual é o assunto mais indexado por ano no repositório.

O Arca é alimentado por 28 comunidades. Cada unidade tem um “núcleo de acesso aberto ao conhecimento” composto por diversos profissionais chaves, dentre os quais existe (na maioria dos casos) um bibliotecário que tem a responsabilidade de preencher todos os metadados corretamente, indicando o assunto devido indexado, verificando o vínculo institucional, e a integridade do objeto digital, fazendo a aprovação final para entrada no repositório.

Em seguida, a equipe de sessão de informação do Arca, responsável pela curadoria dos dados, realiza um processo de verificação de conformidade e em caso de ocorrência de edição, sinaliza para o bibliotecário da unidade alimentadora realizar o ajuste necessário.

Na ferramenta de visualização de dados, a indexação é apresentada em português, mas existem materiais indexados em diversas línguas, como espanhol, inglês, francês e até chinês. Foi criada então uma “tabela de equivalência” que além de tratar a questão de língua, é útil para eliminar sinônimos, homônimos, plurais e variação de grafia. A tabela implementada internamente no sistema através do Kibana e o Elasticsearch, permite que os termos que possam estar oferecendo grafia diversa em distintos documentos, possam ser reunidos na visualização de dados por um único indexador eleito como preferido, mas que ao ser clicado pelo pesquisador, lhe será ofertado todos os resultados indexados com o mesmo termo em diferentes línguas, ou grafados de forma diferente na indexação realizadas pelas unidades depositantes.

Através deste recurso, a indexação realizada pelos bibliotecários de cada unidade é respeitada, mas através da curadoria no Arca, consegue-se ofertar a reunião de todos os conteúdos com base na tabela de equivalências criadas.

#### 4 Análises

Num esforço de sintetizar uma visão geral sobre os SOC que puderam ser conhecidos a partir das entrevistas realizadas, elaborou-se o Quadro 5, com o objetivo de cada SOC e as dificuldades consideradas pelas unidades entrevistadas.

**Quadro 5 - Visão geral sintética dos SOC utilizados.**

SOC Identificados (mesmo se não em uso)	Objetivo	Crítica
DeCS	Indexar a produção de Ciência da Saúde na Fiocruz promovendo sua população no meio da América Latina.	Sem relato de dificuldades no uso. Queixa de demora na atualização dos descritores.

**Quadro 5 - Visão geral sintética dos SOC utilizados.**

<b>SOC Identificados (mesmo se não em uso)</b>	<b>Objetivo</b>	<b>Crítica</b>
CDD	Servir de instrumento para classificar, especialmente os materiais físicos, para localizá-los e organizá-los no acervo conforme disponibilidade física das Unidades de Informação.	Sem relato de dificuldades no uso. Queixa de demora na atualização. O recurso <i>Web</i> deste SOC é entendido como um facilitador para diminuir o tempo de acesso às atualizações de termos, por não depender de aquisição da fonte física.
Lista de Áreas Temáticas da Coleção Infantojuvenil	Facilitar o acesso aos itens do acervo infantojuvenil, com uso de cores.	Sem relato de dificuldades.
Descritores Locais.	Complementar a indexação não coberta pelo DeCS.	Sem relato de dificuldades.
Tesouro de Manguinhos (Teman)	Indexar os materiais submetidos a este recurso de indexação para recuperação a partir dos seus conteúdos, desejando ser mais abrangente que a cobertura do DeCS por considerar outros contextos em que a Ciência de Saúde na Fiocruz faz relação.	Na plataforma em que está implementado, não oferece visualização gráfica da hierarquia retardando a consolidação do nível de especificidades dos conceitos. Relata-se presença de termos construídos com linguagem natural, contrariando os princípios de construção de tesouros. Apresenta-se fora de uso por falta de atualização dos termos, nos quais ainda é possível encontrar termos não mais considerados pela comunidade científica, e ausência dos termos mais atuais.
Tabela Cutter	Compor número de classificação, com notação de autor.	Sem relato de dificuldades.
Vocabulário controlado próprio que remete ao DeCS	Facilitar a busca dos termos autorizados no DeCS.	Sem relato de dificuldades no uso. A queixa está no fato de ser impresso e não contar com uma atualização.
Thesaurus Brasileiro da Educação do INEP	Complementar a indexação não coberta pelo DeCS.	Sem relato de dificuldades no uso. Recurso necessário para enriquecer o processo de indexação.
Vocabulário Controlado da COC	Complementar a indexação não coberta pelo DeCS.	Sem relato de dificuldades.
Tabela de equivalência	Viabilizar o uso da ferramenta de interface definida como "Visualização de Dados", com a função de unificar em "marcador" único a indexação variada por línguas, eliminar sinônimos, homônimos, plurais e variação da grafia.	Sem relato de dificuldades.

Fonte: Os autores (2021).

Apoiados no conteúdo das entrevistas e nos conhecimentos teóricos aprofundados com a pesquisa bibliográfica, seguem análises sobre alguns dos soc evidenciados na exploração. É importante ressaltar que o DeCS possui uma análise mais aprofundada por ser o soc mandatário da Fiocruz e estar disponível para acesso *on-line*. Sobre os demais soc, algumas características foram abordadas a partir das informações obtidas nas entrevistas com os representantes das Unidades de Informação desta instituição.

### a) DeCS - Descritores em Ciências da Saúde

O DeCS foi o Sistema de Organização do Conhecimento identificado como mandatário na instituição. Está disponível *on-line*, por isso foi possível obter informações de algumas características e funcionalidades verificadas em sua página da *web*.

É designado como um vocabulário estruturado, definido em seu próprio *site* como “coleções de termos, organizados segundo uma metodologia na qual é possível especificar as relações entre conceitos com o propósito de facilitar o acesso à informação” (BIREME; OPAS; OMS, [2016?]). Foi elaborado pela BIREME (Centro Latino-Americano e do Caribe de Informação em Ciências da Saúde) no final dos anos 1970 e desenvolvido a partir do MeSH – Medical Subject Headings da U.S. National Library of Medicine (NLM) dos Estados Unidos.

Segundo informações de seu próprio *site*, é um vocabulário multilíngue, com termos em português, inglês, espanhol e francês. Percebe-se que funciona como uma tradução dos termos do MeSH. Colabora também com o projeto de desenvolvimento de terminologia única e rede semântica em saúde, UMLS – Unified Medical Language System da NLM, com envio dos termos MeSH em português e espanhol.

A principal área abrangida é a de Ciências da Saúde com termos médicos originais do MeSH, mas também foram desenvolvidas as áreas específicas de Saúde Pública, Homeopatia, Ciência e Saúde e Vigilância Sanitária.

Foi criado para ser uma linguagem única e padronizada para indexação da literatura técnica e científica das ciências biomédicas e da saúde e recuperação nas bases de dados da Biblioteca Virtual em Saúde (BVS), LILACS, MEDLINE e outras. De modo a facilitar o acesso à informação científica na Região da América Latina e do Caribe (AL&C) com uso de terminologia comum para pesquisa em múltiplos idiomas.

Em seu *site*, há uma breve apresentação do vocabulário na opção “Sobre o DeCS”, onde disponibilizam links para definições de palavras que podem não ser de conhecimento comum, como por exemplo, “vocabulário estruturado”. Contém também *links* que direcionam para bases como o MeSH e Lilacs.

Há também o *link* para a página wiki do DeCS<sup>12</sup>, onde traz mais informações sobre o mesmo. Nesta, possuem instruções de uso na base, mas remete à *prints* de telas com interface diferente da atual, provavelmente por não terem sido atualizadas. Observou-se que alguns *links* desta página também não estão funcionando. Considerou-se importante trazer a informação abaixo retirada desta página:

12 Disponível em: [https://wiki.bireme.org/pt/index.php?title=Descritores\\_em\\_Ci%C3%AAncias\\_da\\_Sa%C3%BAde&redirect=no](https://wiki.bireme.org/pt/index.php?title=Descritores_em_Ci%C3%AAncias_da_Sa%C3%BAde&redirect=no)

O DeCS segue a tradição dos sistemas de classificação e respectivas listas de cabeçalhos de assunto que foram sendo transformadas em vocabulários especializados sem, no entanto, abandonar as estruturas dos sistemas de classificação das quais são originários. Sua estrutura hierárquica é fundamentada na divisão do conhecimento em classes e subclasses decimais respeitando as ligações conceituais e semânticas, e seus termos são apresentados em uma estrutura híbrida de pré e pós coordenação (BIREME; OPAS; OMS, 2020).

A página principal do DeCS oferece opção de visualização nos mesmos idiomas do vocabulário (português, inglês, espanhol e francês), e um menu de acessibilidade. A busca pode ser realizada nos mesmos idiomas e das seguintes formas: “Qualquer termo”, “Termo exato”, “ID do descritor”, “Código hierárquico”, “Qualquer qualificador”. Além do uso dos truncadores \* (asterisco) ou \$ (cifrão) no começo ou no fim do termo de busca.

Verifica-se que os descritores são apresentados nos quatro idiomas e traz a definição em português no campo “Nota de Escopo”. A “Nota de indexação” traz informações de uso dos descritores. Há ainda outros campos que não foram abordados, pois a intenção deste trabalho não é fazer uma análise exaustiva. Apresenta também um índice alfabético, porém, é difícil de ser encontrado dentro do *site* (identificado na opção “Edição atual”).

Foram encontradas as relações de equivalência em “Termos alternativos” (Remissivas - USE, UP) e relações hierárquicas na aba “Estrutura hierárquica” (TG - Termo Geral e TE - Termo Específico). Já em “Veja também os descritores” acredita-se que corresponda às relações associativas (TR - Termo Relacionado, Remissiva Ver Também). Porém, essa informação não está clara, assim, não é seguro fazer esta afirmação, já que há algumas inconsistências entre os *links* que direcionam aos descritores.

As relações hierárquicas, apesar de não descritas claramente, podem ser visualizadas facilmente em formato de árvore na aba “Estrutura hierárquica”. Como exemplo temos: o termo “Ocupações em Saúde” é um TG de “Medicina”, o qual está no mesmo nível de “Acupuntura”, por exemplo. E o termo “Medicina do Vício”, assim como os demais abaixo de “Medicina”, são compreendidos como seu TE.

Foram identificados avisos sobre descritores novos em português, inglês e espanhol, descritores alterados com suas respectivas substituições, e descritores excluídos com indicação para o conceito absorvido, o que é importante para seu usuário saber das mudanças ocorridas.

Ademais, oferece informações sobre mudanças de códigos hierárquicos DeCS e MeSH, dados estatísticos (Gráficos: Distribuição de novos conceitos por categoria, Variação na quantidade de descritores e sinônimos por idioma, Distribuição dos conceitos por categoria), visão geral com quantitativo de descritores novos por categoria e subcategoria, quantitativo de mudanças realizadas no ano, revisões e fusões realizadas nos descritores e ainda disponibilização de edições anteriores do vocabulário (desde 1999). Foi interessante notar também a disponibilidade para sugestões de novos termos, por meio de preenchimento de formulário, além de contato pelo “Fale conosco”.

No geral, é uma ferramenta bem desenvolvida, com muitos recursos e informações importantes, que podem ajudar no entendimento e uso de seu *site* e descritores. Acredita-se atender bem a temática de saúde, visto que é uma linguagem muito usada pela área no Brasil. No que tange às unidades entrevistadas da Fiocruz, a maior dificuldade apontada é a demora na atualização dos termos. No entanto, é uma ferramenta satisfatória de uso geral por toda rede de bibliotecas.

#### **b) Classificação Decimal de Dewey (CDD)**

A Classificação Decimal de Dewey (CDD) tem caráter universalista com um grau de especificidade que atende as necessidades da instituição. Utilizada na maioria das unidades em sua versão impressa, é considerada pela comunidade como suficiente e apropriada.

Algumas unidades usuárias deste SOC percebem insuficiências de especificidades em alguns poucos casos, ou lentidão na atualização. A proposta de acesso ao *Web Dewey* (versão *on-line*) para uso de toda Rede de Bibliotecas parece uma solução muito apropriada para eliminar a desatualização que a versão impressa impõe por seu alto custo de aquisição.

#### **c) Tesouro de Manguinhos (Teman)**

Conforme relato dos profissionais entrevistados, o Tesouro de Manguinhos (Teman) foi criado em 1986, a partir de projetos financiados pela FINEP, tendo como objetivo subsidiar o processo de indexação da Biblioteca de Manguinhos. Foi muito utilizado para indexar o projeto Produção de Manguinhos (PROMAN), que reúne toda produção científica dos pesquisadores da instituição. No entanto, encontra-se fora de uso em função de sua desatualização.

Segundo Tálamo, Lara e Kobashi (1992) o tesouro é

[...] um objeto **cultural** que registra e representa o conhecimento segundo parâmetros estáveis e previamente determinados. Tais parâme-

tros, que se materializam sob a forma de **redes de relações entre descritores**, determinam [...] o modo de organização e disseminação da informação [...] (TÁLAMO; LARA; KOBASHI, 1992, p. 198, grifos nossos).

Com base nessa proposição, o Teman qualifica-se como um tesouro por representar o conhecimento conforme a cultura da instituição, através de uma rede de relações entre descritores estruturados em forma de árvore terminológica. Sua única divergência é não ter implementado um recurso gráfico para oferecer a visualização dessa estrutura hierárquica. A estrutura existe, haja vista, a identificação de termos gerais, termos específicos, termos relacionados e associativos presentes dentro do sistema.

Até mesmo os conceitos podem ser recuperados dentro do sistema, notadamente uma característica dos tesouros. Mas uma visualização de sua árvore não é oferecida ao indexador para que o mesmo possa ter uma noção durante a indexação, de quão específico ou geral o termo pesquisado é dentro da estrutura. Este conhecimento estaria mais próximo de profissionais já experientes no processo de indexação da unidade, por força da vivência dentro da temática. A ausência da visualização de sua árvore hierárquica pode ser um fator que retarda a adaptação de novos profissionais às atividades de indexação, especialmente quando este não trazer uma bagagem de conhecimento sobre a temática, quando de sua admissão.

Entendido como um recurso rico para representar as particularidades da instituição, lamentamos seu status de “desatualizado” e “fora de uso”. A atualização de sua estrutura é um projeto que está sendo desenhado por um grupo de trabalho composto por profissionais da instituição.

Acreditamos que, se seu escopo de cobertura transcender os limites das áreas especializadas da Biblioteca de Manguinhos e do projeto PROMAN, e se dispuserem a abarcar todos os campos de atuação da Fiocruz, ter-se-ia um SOC robusto e capaz de representar todo o conhecimento da instituição e conectar toda gama de acervos (ao menos internamente).

#### **d) Vocabulário Controlado da COC**

O Vocabulário Controlado da Biblioteca Virtual da Casa de Oswaldo Cruz (COC) foi desenvolvido frente a necessidade de promover a “conversa” entre diferentes acervos (arquitetônico, arquivístico, bibliográfico, museológico e virtual). Consiste em uma lista alfabética de assuntos com os termos mais empregados em todos os acervos da COC. É chamado de vocabulário controlado pois a intenção é que se torne um, mas ainda se configura como uma lista de cabeçalho de assuntos.

Mostrou-se como solução inovadora para conectar em uma mesma plataforma, itens de diversas tipologias documentais. Acervos distintos que têm relações contextuais entre si agora podem ser acessados por um recurso único.

Sua implementação ainda é rudimentar por estar estruturada em planilha à parte ao recurso de conexão, ou seja, à biblioteca virtual. Entendemos que sua implementação em meio *web* seria extremamente vantajoso para futuros projetos de interoperabilidade.

#### **e) Classificação Infantojuvenil**

Inspirada na classificação por cores, comumente empregada para distinguir as faixas etárias a qual se destina a obra, a Lista de Áreas Temáticas da Coleção Infantojuvenil tem a proposição de evidenciar as temáticas trabalhadas pela unidade (Museu da Vida) vinculada à Biblioteca de Educação e Divulgação Científica Iloni Seibel.

O recurso de cor promove grande vantagem no agrupamento dos itens temáticos para os preparadores de atividades e conteúdos, assim como para os usuários que frequentam a unidade com o propósito de explorá-los.

### **5 Considerações finais**

Entendemos que estudos semelhantes a este, no qual é permitido registrar as deficiências identificadas, são úteis como recurso para obter conhecimento, evidenciar a necessidade de aprimoramentos e servir de suporte para o planejamento de projetos ou ações orientadas ao desenvolvimento.

Neste sentido, inferimos que este trabalho foi útil ao seu propósito acadêmico, sem objetivar gerar qualquer ônus crítico à instituição. Na contrapartida a toda benevolente disponibilidade apresentada por todos os profissionais entrevistados, o trabalho encontra-se disponível para o uso da instituição na forma que lhe seja útil.

Percebeu-se que embora o estudo tenha sido realizado sob a particularidade de uma única instituição, se há configuração constituinte de diversas Unidades de Informação focadas em grupos de áreas, por vezes distintos, torna-se necessário um estudo sobre os SOC que melhor atenderão cada unidade.

É notável que alguns SOC têm proposição de agrupar e representar a produção nas diversas atuações (CDD e/ou DeCS), mas também, que nem sempre é possível atender a todas as particularidades com um mesmo SOC. Nestes contextos, é inegável a necessidade de recorrer a outras soluções que possam enriquecer a noção de capacidade de atuação da instituição: Teman, “Vocabulário Controlado” da COC, Classificação temática por cores para Coleção Infantojuvenil.



Foi enriquecedor perceber o engajamento de profissionais focados na obtenção destes sucessos, criando e implementando soluções que atendam suas unidades e consequentemente engrandecem o poder da instituição.

A autonomia das unidades, sem negar as propostas de conexão geral, foi observada como combustível para gerar liberdade de inovações nas soluções para a Organização do Conhecimento.

Assim, as Unidades de Informação da Fiocruz seguem cumprindo seu papel de organizar a informação da instituição e promover o acesso aos seus usuários, com auxílio das ferramentas apresentadas.

## Referências

BARDIN, Laurence. **Análise de conteúdo**. São Paulo: Edições 70, 1977.

BIREME; OPAS; OMS. **DeCS - Descritores em Ciências da Saúde**. [São Paulo], 2016. Página *html* do DeCS. Disponível em: <http://decs2016.bvsalud.org/P/decsweb2016.htm>. Acesso em: 22 set. 2020. (a)

BIREME; OPAS; OMS. Descritores em Ciências da Saúde. In: BIREME; OPAS; OMS. **Wiki BIREME**. [S. l.], 2020. Página *html* tópica na Wiki BIREME. Disponível em: [https://wiki.bireme.org/pt/index.php?title=Descritores\\_em\\_Ci%C3%A2ncias\\_da\\_Sa%C3%BAde&redirect=no](https://wiki.bireme.org/pt/index.php?title=Descritores_em_Ci%C3%A2ncias_da_Sa%C3%BAde&redirect=no). Acesso em: 22 set. 2020.

BIREME; OPAS; OMS. Vocabulário estruturado. BIREME; OPAS; OMS. **DeCS - Descritores em Ciências da Saúde**. [São Paulo, 2016?]. Página *html* do DeCS sobre o link vocabulário estruturado. Disponível em: <https://decs.bvsalud.org/P/aboutvocabp.htm>. Acesso em: 22 set. 2020. (b)

CARLAN, Eliana; MEDEIROS, Marisa Bräscher Basílio. Sistemas de Organização do Conhecimento na visão da Ciência da Informação. **Revista Ibero-americana de Ciência da Informação**, Brasília, v. 4, n. 2, p. 53-73, ago./dez. 2011.

DAHLBERG, Ingetraut. Knowledge organization: a new science? **Knowledge Organization**, Baden-Baden, v. 33, n. 1, p. 11-19, 2006.

DESCRITORES em Ciências da Saúde: DeCS. Ed. 2020. São Paulo: BIREME; OPAS; OMS, 2020. Disponível em: <http://decs.bvsalud.org>. Acesso em 22 set. 2020.

FUNDAÇÃO OSWALDO CRUZ. **Bibliotecas**. Rio de Janeiro, [201-?]. Página *html* da Fiocruz. Disponível em: <https://portal.fiocruz.br/bibliotecas>. Acesso em: 10 set. 2020. (a)

FUNDAÇÃO OSWALDO CRUZ. **Bibliotecas físicas**. Rio de Janeiro, [201-?]. Página *html* da Fiocruz. Disponível em: <https://portal.fiocruz.br/bibliotecas-fisicas>. Acesso em: 10 set. 2020. (b)

FUNDAÇÃO OSWALDO CRUZ. **Bibliotecas virtuais em saúde**. Rio de Janeiro, [201-?]. Página *html* da Fiocruz. Disponível em: <https://portal.fiocruz.br/bibliotecas-virtuais-em-saude>. Acesso em: 10 set. 2020. (c)

FUNDAÇÃO OSWALDO CRUZ. Casa de Oswaldo Cruz. **A biblioteca virtual Oswaldo Cruz**. Rio de Janeiro, [201-?]. Página *html* da BVOC. Disponível em: <http://oswaldocruz.fiocruz.br/index.php/sobre-a-biblioteca-virtual-oswaldo-cruz>. Acesso em: 10 set. 2020. (d)

FUNDAÇÃO OSWALDO CRUZ. **Escola Politécnica de Saúde Joaquim Venâncio**. Rio de Janeiro, [201-?]. Página *html* da Fiocruz. Disponível em: <http://www.epsjv.fiocruz.br/>. Acesso em: 16 out. 2020.

FUNDAÇÃO OSWALDO CRUZ. Instituto de Comunicação e Informação Científica e Tecnológica em Saúde. **Bibliotecas**. Rio de Janeiro, [201-?]. Página *html* do ICICT. Disponível em: <https://www.icict.fiocruz.br/bibliotecas>. Acesso em: 10 set. 2020. (d)

FUNDAÇÃO OSWALDO CRUZ. **Teses e dissertações**. Rio de Janeiro, [201-?]. Página *html* da Fiocruz. Disponível em: <https://portal.fiocruz.br/teses-e-dissertacoes-o>. Acesso em: 10 set. 2020.

FUNDAÇÃO OSWALDO CRUZ. **Teses Fiocruz**. Rio de Janeiro, [2018?]. Base de dados da Fiocruz. Disponível em: <http://teses.icict.fiocruz.br>. Acesso em: 10 set. 2020..

KAUARK, Fabiana; MANHÃES, Fernanda Castro; MEDEIROS, Carlos Henrique. **Metodologia da pesquisa: guia prático**. Itabuna: Via Litterarum, 2010.

LEONARDO. Sobre a Biblioteca: missão. FUNDAÇÃO OSWALDO CRUZ. Instituto de Comunicação e Informação Científica e Tecnológica em Saúde. **Biblioteca da Saúde da Mulher e da Criança**. Rio de Janeiro, 2008. Página *html* da Biblioteca da Saúde da Mulher e da Criança. Disponível em: <http://www.fiocruz.br/bibsmc/cgi/cgilua.exe/sys/start.htm?infoid=53&sid=2>. Acesso em: 30 out. 2020.

MAGALHÃES, Rosicler. Sobre a Biblioteca: missão. FUNDAÇÃO OSWALDO CRUZ. Instituto de Comunicação e Informação Científica e Tecnológica em Saúde.

**Biblioteca de Manguinhos.** Rio de Janeiro, 2017. Página *html* da Biblioteca de Manguinhos. Disponível em: <http://www.fiocruz.br/bibmang/cgi/cgilua.exe/sys/start.htm?sid=77>. Acesso em: 10 set. 2020.

MELO, Maria Antônia F.; BRÄSCHER, Marisa. Termo, conceito e relações conceituais: um estudo das propostas de Dahlberg e Hjørland. **Ci. Inf.**, Brasília, DF, v. 41 n. 1, p.67-80, jan./abr., 2014.

MINAYO, M. C. S. **O desafio do conhecimento:** pesquisa qualitativa em saúde. São Paulo: Rio de Janeiro: HUCITEC/ABRASCO, 1992.

MORAES, I. S.; MOREIRA, W. Os conceitos de sistemas de organização do conhecimento e linguagens documentárias: análise de domínio. **Encontro Nacional de Pesquisa em Ciência da Informação**, n. XIX ENANCIB, 2018. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/103667>. Acesso em: 17 nov. 2020.

SOUZA, Rosali Fernandez de. Organização do conhecimento. In.: TOUTAIN, Lídia M. B. B. (org). **Para entender a Ciência da Informação.** Salvador: EDUFBA, 2007. p. 103-124.

TÁLAMO, Maria de Fátima G. Moreira; LARA, Marilda Lopes Ginez de; KOBASHI, Nair Yumiko. Contribuição da terminologia para a elaboração de tesouros. **Ciência da Informação**, Brasília, v. 21, n. 3, p. 197-200, set./dez. 1992.

## Apêndice A - Roteiro de entrevista

### Contextualização:

Esta entrevista faz parte da metodologia adotada pelos entrevistadores para um levantamento quanti-qualitativo com intuito de investigar os Sistemas de Organização do Conhecimento (soc) utilizados na Fundação Oswaldo Cruz (Fiocruz).

Em virtude do contexto atual de pandemia, esta entrevista será realizada de forma *on-line*, via plataforma *Google Meet*. O objetivo pretendido com uso dos questionamentos abaixo relacionados é identificar, dentro do escopo do campo empírico selecionado: Rede de Bibliotecas da Fiocruz, a existência de Sistemas de Organização do Conhecimento e a verificação de seus desempenhos.

Sistemas de Organização do Conhecimento (soc) - São sistemas conceituais semanticamente estruturados que contemplam termos, definições, relacionamentos e propriedades dos conceitos. Na organização e recuperação da informação, os soc cumprem o objetivo de padronização terminológica para facilitar e orientar a indexação e os usuários. Abrangem todos os tipos de esquemas que organizam e representam o conhecimento, por exemplo: classificações, taxonomias, tesouros, ontologias (CARLAN; MEDEIROS, 2011, p. 54).

### Entrevistadores:

- Deniz Costa (Universidade Federal do Estado do Rio de Janeiro - UNIRIO)
- Elaine Lucia (Instituto Brasileiro de Informação em Ciência e Tecnologia - IBICT/UFRJ)
- Karina Veras (Instituto Brasileiro de Informação em Ciência e Tecnologia - IBICT/UFRJ)
- Thaís Sant'Anna (Universidade Federal do Estado do Rio de Janeiro - UNIRIO)

### Cadastro do respondente:

– Aspectos gerais

- a) Responsável entrevistado:
- b) Contato:
- c) Nome da Biblioteca:
- d) Local:
- e) Equipe:

- f) Domínio de conhecimento coberto pela biblioteca:
- g) Público-alvo da biblioteca:

**Roteiro:**

– Aspectos específicos ao SOC:

- h) Utilizam algum Sistema de Organização do Conhecimento (SOC)? Se sim, qual ou quais?
- i) Como é sua apresentação: física ou digital? Se digital é possível termos acesso ao mesmo?
- j) Se existir mais de 1, eles possuem relação?
- k) Pode falar um pouco sobre a história do(s) SOC em sua biblioteca?
- l) Seu uso é restrito a essa biblioteca ou é de aplicação geral/institucional?
- m) Considera existir dificuldades na aplicação desse(s) SOC? Fale-nos sobre.
- n) Se sim, como entende que poderia ser aprimorado?
- o) Como analisa o desempenho da recuperação da informação com uso do(s) SOC relacionado(s)?
- p) Conte-nos sobre projetos e ações orientadas para o desenvolvimento de SOC em sua instituição.



# Os sistemas de organização do conhecimento das universidades: um estudo de caso das bibliotecas da Universidade Federal do Rio de Janeiro

Carla Beatriz Marques Felipe<sup>1</sup>, Eliane Azevedo Gomes<sup>2</sup>, Milene Couras da Cruz<sup>3</sup>

---

## 1 Introdução

OS SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO (SOC) SÃO INSTRUMENTOS utilizados tanto na organização quanto na gestão do conhecimento, visando a recuperação da informação, mais especificamente a sua representação. As bibliotecas universitárias fornecem informações para toda a comunidade acadêmica como estudantes, professores, técnicos administrativos e pesquisadores, fomentando as ações dos três pilares da universidade: ensino, pesquisa e extensão.

A presente pesquisa se caracteriza como uma pesquisa exploratória e descritiva, com abordagem quali-quantitativa. Este capítulo fará um estudo de caso da Universidade Federal do Rio de Janeiro (UFRJ), criada em 1920, pelo governo federal a partir da união da Escola Politécnica (1792), da Faculdade Nacional de Medicina (1808) e da Faculdade Nacional de Direito (1891). A UFRJ é um espaço onde a informação é um fator primordial, visto que se dedica a formação acadêmica de diferentes áreas do conhecimento, associado à atuação relevante na comunidade como um todo e nos aspectos inovadores das pesquisas que impactam na sociedade. A estrutura reúne seis centros com as respectivas unidades e quatro campi avançados. As 44 bibliotecas são coordenadas pelo Sistema de Bibliotecas e Informação (SiBI).

---

1 Doutoranda em Ciência da Informação - PPGCI- IBICT/UFRJ; Mestre em Ciência da Informação - PPGCI-UFPE; Universidade Federal do Rio de Janeiro - Departamento de Biblioteconomia e Gestão de Unidade de Informação. felipecarla12@gmail.com

2 Doutoranda em Ciência da Informação - PPGCI- IBICT/UFRJ ; Mestre em Biblioteconomia- PPG-UNIRIO; Universidade Federal do Rio de Janeiro. elianeagomes@gmail.com

3 Mestranda em Ciência da Informação - PPGCI- IBICT/UFRJ; Especialista em Comunicação Integrada - ESPM/RJ; Pontifícia Universidade Católica do Rio de Janeiro – Centro de Ciências Biológicas e da Saúde. milene.couras@gmail.com

O SiBI foi criado em 1983 como órgão suplementar do Fórum de Ciência e Cultura (FCC) com a função de gerenciar as bibliotecas. Tem como objetivo dar suporte às questões de ensino, pesquisa e extensão da instituição, além de integrar as bibliotecas na política educacional e administrativa da Universidade. O SiBI promove o desenvolvimento das bibliotecas a partir da capacitação continuada de seus membros, da modernização e informatização, da atualização e manutenção dos acervos, da definição de políticas de informação e de padrões técnicos. As bibliotecas que compõem o SiBI estão organizadas em oito centros, a saber: Centro de Ciências Jurídicas e Econômicas (CCJE), Centro de Ciências Matemáticas e da Natureza (CCMN), Centro de Ciências da Saúde (CCS), Centro de Filosofia e Ciências Humanas (CFCH), Centro de Letras e Artes (CLA), Centro de Tecnologia (CT), Fórum de Ciência e Cultura (FCC), Macaé Campus Professor Aloísio Teixeira.

O objetivo desta pesquisa é investigar quais são os sistemas de organização do conhecimento utilizados pelas Bibliotecas que compõem o SiBI da UFRJ. Os objetivos específicos são: identificar as áreas do conhecimento das bibliotecas; mostrar quais os tipos de linguagens documentárias são utilizados pelas bibliotecas; e descrever as características das linguagens documentárias utilizadas pelas bibliotecas fazendo um paralelo com a literatura da área de organização do conhecimento.

Como instrumento de coleta de dados foi elaborado um questionário (apêndice A), no *google forms*, dividido em três seções. A primeira seção foi dividida em duas partes. A primeira parte voltada para a identificação das bibliotecas respondentes com três perguntas que visavam saber o nome da biblioteca, o centro ao qual está vinculada e as áreas do conhecimento que atende. A segunda parte está destinada a identificar se essas bibliotecas possuem um controle de vocabulário construído, adaptado ou apropriado pela biblioteca e qual seria esse vocabulário. Foi entendido como instrumento construído aquele que foi elaborado pela equipe da biblioteca anterior ou atual. Instrumentos que foram apropriados são aqueles que foram incorporados como único instrumento utilizado pela biblioteca. Instrumentos adaptados são aqueles que já existem, mas foram adaptados para a realidade da biblioteca. A segunda seção do questionário é orientada àqueles que construíram um instrumento próprio, buscou identificar o propósito, os profissionais envolvidos, o tempo destinado a sua construção, o processo de coleta e definição dos termos, assim como sua atualização, a forma de manutenção e o idioma. A terceira seção do questionário dirigida àqueles que não construíram seus instrumentos, mas utilizam algum tipo de controle de vocabulário, as perguntas foram voltadas para identificar qual o instrumento, o seu formato, sua atualização, seu idioma, se a biblioteca sente a necessidade de construir um instrumento próprio, por qual motivo e se pretende construir.



O questionário foi enviado por e-mail à Coordenação do SiBI em 11 de janeiro de 2021, solicitando a anuência e a cooperação da coordenação ao acesso às bibliotecas do sistema. A solicitação foi prontamente atendida pela coordenadora do SiBI, que encaminhou o e-mail para a Divisão de Processamento Técnico que os encaminhou as bibliotecárias responsáveis pelas equipes de catalogação de cada biblioteca. O maior desafio desta pesquisa foi conseguir que as 44 bibliotecas respondessem ao questionário. Para tentar solucionar este obstáculo foram enviados sucessivos e-mails para as bibliotecas e suas chefias. Em 28 de janeiro de 2021 foram enviados e-mails diretamente às bibliotecas e suas chefias. No dia 04 de fevereiro de 2021 foram enviados e-mails às equipes de catalogadores que foram possíveis identificar nos sites das bibliotecas. E no dia 17 de fevereiro de 2021 foram re-enviados e-mails diretamente às bibliotecas e suas chefias. As bibliotecas também foram contactadas via mídias sociais. Após todas essas estratégias, 32 bibliotecas responderam ao questionário.

A etapa seguinte previa a compilação dos dados adquiridos a partir das respostas dos questionários a fim de identificar os Sistemas de Organização do Conhecimento utilizados para o controle do terminológico das bibliotecas que compõem o sistema. Os dados foram tabulados em planilha no software *excel* permitindo sua análise. Após o mapeamento e a tabulação dos SOC adotados pelas bibliotecas, estes foram analisados individualmente buscando identificar as características de cada SOC, sua temática e estrutura. Para esta análise, os SOC foram visitados em seus *websites*.

O capítulo está dividido em uma parte teórica onde são explanados os conceitos de SOC e seus tipos; os resultados da aplicação dos questionários; e uma discussão trazendo os SOC utilizados pelas bibliotecas que compõem o SiBI e a literatura científica existente na área de Organização do Conhecimento.

## **2 Sistemas de Organização Conhecimento**

O conhecimento ocorre na estrutura cognitiva do indivíduo por meio do processamento de informações recebidas no dia a dia. Pode-se dizer que é difícil transmissão e está associado à percepção e ao aprendizado do mundo. Segundo Felipe (2016, p. 57), “é por meio do conhecimento que se tem consciência de como se deve agir e pensar com relação à própria vida”. O conhecimento fornece ao indivíduo as informações necessárias para a tomada de decisão no que diz respeito ao trabalho, ciência, relacionamentos, família e até mesmo a religião. Sendo parte importante para os mais variados aspectos da vida do indivíduo, pode-se afirmar que o conhecimento necessita ser organizado para ser disseminado e utilizado.

Nesse sentido, a Organização do Conhecimento (OC) é primordial para a sociedade, pois se ocupa em tornar o conhecimento disponível. A OC pode ser entendida como a disciplina que trata sobre as propriedades do conhecimento, sobretudo no que diz respeito à recuperação da informação, utilizando-se de instrumentos para sua representação tendo como foco o conhecimento registrado.

Conforme Nascimento e Pinho (2019, p. 107), a “OC objetiva a sistematização e organização do conhecimento com a finalidade de recuperação da informação, considerada também como um campo que exerce um papel de comunicação, mediação e transferência de conhecimento [...]”. Portanto, sem o que é desenvolvido na OC, o conhecimento não pode ser recuperado, logo a sua utilização não seria possível. Segundo Hjørland (2008, p.86, tradução nossa), a Organização do Conhecimento:

como um campo de estudo está preocupado com a natureza e qualidade de tais processos de organização do conhecimento [...] bem como os sistemas de organização do conhecimento [...] usado para organizar documentos, representações de documentos, obras e conceitos.

Quando Hjørland disserta sobre os processos de OC, ele se refere aos processos de indexação, classificação e representação da informação objetos da Biblioteconomia e Ciência da Informação. Por sua vez, Smiraglia (2012, p. 225, tradução nossa) aponta que “a Organização do Conhecimento (também conhecida pela sigla KO, do inglês) é o domínio onde o ordenamento do conhecimento é o paradigma principal de investigação científica, cuja aplicação básica é o desenvolvimento de sistemas”. Isso significa que a organização, a padronização e o tratamento dado ao conhecimento é o foco central da Organização do Conhecimento, no desenvolvimento de teorias e na elaboração de instrumentos e sistemas que facilitam essa ordenação.

Nesse sentido, faz uso dos Sistemas de Organização do Conhecimento (SOC) que são mecanismos que auxiliam as atividades desenvolvidas pela Organização do Conhecimento, a fim de representar o conhecimento. Segundo Carlan e Bräscher (2011, p. 55):

No contexto da CI, os SOCs ou esquemas de representação do conhecimento são instrumentos que fazem a tradução dos conteúdos dos documentos originais e completos, para a um esquema estruturado sistematicamente, que representa esse conteúdo, com a finalidade principal de organizar a informação e conhecimento e, conseqüentemente, facilitar a recuperação das informações contidas nos documentos.

Os soc são os instrumentos conhecidos anteriormente como linguagens documentárias, mas que com o advento das tecnologias englobam elementos de inovação e novos princípios, como ontologia, que passaram a auxiliar a Organização do Conhecimento. Segundo Fujita, Santos e Alves (2018, p. 24) “na literatura da área de Organização do Conhecimento, essas linguagens ou sistemas são tipos de instrumentos cujos conjuntos são chamados de “linguagem de indexação”, “linguagem documentária”, e “Sistemas de Organização do Conhecimento[...]”. Ainda segundo os autores, a adoção do termo Sistemas de Organização do Conhecimento é novo na área de Organização do Conhecimento.

Para Barité (2011, p. 126), “o objetivo principal de cada uma dessas línguas [soc] continua a ser a representação temática do conteúdo de qualquer documento, recurso pacote de informação”. Ou seja, mesmo com a mudança na terminologia, e acréscimo de novos recursos, o propósito dos soc continua o mesmo, auxiliar no momento da tradução dos termos da linguagem natural para a linguagem documentária.

Antes de explicar o funcionamento de cada um dos sistemas, é preciso entender um pouco sobre o conceito que é base para o progresso da OC e auxilia no entendimento do funcionamento e desenvolvimento dos soc. Para Bräscher e Carlan, os conceitos são os elementos essenciais na elaboração de um soc, uma vez que “são as unidades básicas do conhecimento” (BRÄSCHER; CARLAN, 2010, p. 155).

Segundo Dahlberg (1978), o conceito é constituído de enunciados que apresentam a característica de uma determinada coisa, objeto, usando a linguagem natural. Essas características tornam o objeto ou a sua definição, aqui entendida como conceito de único, sem igual. Assim, o conceito de SISTEMAS DE BIBLIOTECAS E INFORMAÇÃO DA UFRJ pode ser constituído com os seguintes enunciados:

- a) É constituído por 44 bibliotecas;
- b) Foi fundado em 1983;
- c) Estão ligados a UFRJ;
- d) Estão no Estado do Rio de Janeiro.

Os enunciados aqui apresentados são verdadeiros e unidos constituem o conceito de SISTEMAS DE BIBLIOTECAS E INFORMAÇÃO DA UFRJ. Para Dahlberg (1978, p. 102), “cada enunciado verdadeiro representa um elemento de conceito”. Isso só é possível por causa da linguagem natural.

São exemplos de soc os Tesouros, Vocabulários Controlados, Sistemas de Classificação Bibliográficas, Ontologia, Taxonomia e a Folksonomia. Cada um desses sistemas surge em uma determinada época, com características e objetivos especí-

ficos. Porém, todos têm o mesmo objetivo que é oferecer suporte a Organização do Conhecimento com a finalidade de favorecer o seu desenvolvimento e a sua gestão.

Com relação à tipologia dos SOC a partir da complexidade das estruturas e principais funções, Zeng (2008) os divide em quatro categorias: 1) lista de termos que inclui as listas propriamente ditas, os dicionários, os glossários e os anéis de sinônimos; 2) modelos semelhantes a metadados que engloba arquivos de autoridades, diretórios e Gazetteers<sup>4</sup>; 3) classificação e categorização que abarca os cabeçalhos de assunto, os esquemas de categorização, as taxonomias e os esquemas de classificação; 4) modelos de relações que abrangem os tesouros, as redes semânticas e as ontologias.

Será apontada a definição dos SOC que apresentam funções com maior pertinência para o uso em bibliotecas universitárias, o que permitirá um melhor entendimento dos resultados da pesquisa.

## 2.1 Tesouro

O termo tesouro vem do Grego e do Latim e significa tesouro. É um dos SOC mais antigos e podem ser tanto temáticos, quanto multidisciplinares. Como exemplo de tesouros temos o Tesouro Brasileiro de Ciência da Informação, o Tesouro de Folclore e Cultura Popular Brasileira e o Thesaurus Brasileiro da Educação do INEP.

Segundo Moreira González (2011, p. 63), tesouro é uma “lista de descritores (termos controlados) que representam os conceitos de um domínio do conhecimento. E se organiza em estrutura hierárquica com relações semânticas entre si”. Dessa forma, ao apresentar um termo, as formas de uso são exemplificadas na sua própria descrição, levando em consideração o vínculo que o termo tem com os outros termos e a área do conhecimento. Ainda conforme Moreira González (2011, p. 64) “o principal objetivo dos tesouros é eliminar a ambiguidade da linguagem na matéria à qual pertencem os termos”. Isso fica demonstrado na própria estrutura do tesouro, sobretudo quando se trata de tesouros temáticos.

Corroborando com Moreira González, Clarke (2019) disserta que, “a principal função de um tesouro é apoiar a recuperação de informações, orientando a escolha de termos para indexação e pesquisa”. Essa orientação parte da própria estrutura do tesouro, uma vez que os termos são apresentados em ordem alfabética e são apresentadas as relações conceituais de cada termo. Exemplificando as relações, temos na figura abaixo um recorte da apresentação de um conceito e como deve ser usado pelo indexador no Tesouro Brasileiro de Ciência da Informação:

---

4 “Dicionários geoespaciais de lugares nomeados e digitado” (ZENG, 2008)

**Figura 1: Exemplificação das relações conceituais no Tesouro Brasileiro de Ciência da Informação**

<b>tesauros</b>	
ING:	thesauri
ESP:	tesauros
TG	linguagens documentárias
TE	microtesauros
	tesauros multilíngues
TR	controle de vocabulário
	descritores
	dicionários
	elaboração de linguagens documentárias
	listas de cabeçalhos de assunto
	manutenção de tesauros
	notas explicativas
	relações associativas
	relações de equivalência
	sistemas de classificação
	taxonomias
	termos candidatos a descritor
	tesauros facetados
CAT:	2.1.2 Sistemas de organização do conhecimento

Fonte: Tesouro Brasileiro de Ciência da Informação (2014, p. 221).

A figura acima demonstra a relação conceitual de um determinado termo dentro de um tesouro. As siglas TG e TE dizem respeito às relações hierárquicas onde TG é Termo Genérico e TE, Termo Específico. Essa ordenação facilita na escolha do termo por parte do indexador e fica claro como e onde ele deve usar o termo dentro do seu sistema de recuperação da informação. Conforme Tálamo, Lara e Kobashi (1992, p. 198)

o tesouro é um objeto cultural que registra e representa o conhecimento segundo parâmetros estáveis e previamente determinados. Tais parâmetros, que se materializam sob a forma de redes de relações entre descritores, determinam não só o modo de organização e disseminação da informação.

Nesse contexto, utilizando a figura 1 como exemplo, acerca do termo tesauros e suas relações, este é arranjado de acordo com o conhecimento existente e objeto cultural dentro das áreas da Biblioteconomia e da Ciência da Informação. Nas quais o termo tesouro é uma linguagem documentária, e está totalmente ligado a termos como descritores, controle de vocabulários, relações associativas e outros termos ligados à recuperação da informação, essas relações são conhecidas e utilizadas dentro da Biblioteconomia e da Ciência da Informação de forma natural pelos seus pesquisadores.

## 2.2 Listas de Cabeçalhos de assunto

Ao contrário dos Tesauros, os cabeçalhos de assuntos não são elaborados especificamente para uma determinada área do conhecimento, podem sim, ser so-

mente de uma área do conhecimento, mas em sua maioria abrangem as diversas áreas, apresentando diversos conceitos de maneira mais genérica. O início da sua utilização se deu nos Estados Unidos com a concepção de catálogos alfabéticos pela *Library of Congress* (TÔRRES, 2021) e até o hodierno contexto são estruturados de maneira alfabética.

São exemplos de Cabeçalhos de assuntos a *Library of Congress Subject Headings* (LCSH), organizada pela *Library of Congress*, sendo este habitualmente o mais utilizado nas bibliotecas do Estados Unidos (Harpring, 2016), Terminologia de assuntos da Biblioteca Nacional do Brasil e Descritores em Ciências da Saúde (DeCS). Conforme Cesarino e Pinto (1978, p.273), cabeçalhos de assunto podem ser definidos como “palavra ou grupo de palavras que expressam o conteúdo de um documento”. Ou ainda de vários documentos inclusos nas áreas do conhecimento na qual as bibliotecas estão inseridas. Harpring (2016, p.44) disserta que “listas de cabeçalhos de assunto incluem geralmente listagens separadas de subcabeçalhos padronizados (por exemplo, localizações geográficas), que podem ser combinadas com cabeçalhos designados de acordo com regras preestabelecidas”. A sua estruturação e organização será estabelecida de acordo com a área do conhecimento na qual está designada para representar.

Segundo Moreiro Gonzáles (2011, p.43), os cabeçalhos de assunto apresentam as seguintes características:

Termos controlados de cobertura temática genérica;

Linguagem pré-coordenada: regras para combinar termos no momento da indexação em uma ordem previamente estabelecida.

A recuperação se realiza sequencialmente, pela ordem do termo.

Não é possível recuperar documentos a partir de aspectos não definidos explicitamente nos cabeçalhos.

Quando o autor fala em linguagem pré-coordenada, ele quer dizer que o termo é combinado no momento da indexação pelo indexador. Essa mesma característica é encontrada nos Sistemas de classificação bibliográfica. Diferente dos Tesouros, onde um termo pode ser aplicado no mais variado contexto, nos cabeçalhos de assunto, um termo representa apenas um ponto de acesso. Por isso que a indexação elaborada por meio dos cabeçalhos possui uma ordem pré-estabelecida.

### 2.3 Esquemas de Classificação Bibliográfica

Os esquemas de classificação bibliográficos, como a própria nomenclatura sugere, surgem para auxiliar a organização dos acervos das bibliotecas nas estantes. Porém, além dessa função inicial, os sistemas também são utilizados para organização do acervo por assunto e na parte da indexação. Para Souza (2007, p. 107) estes são os sistemas mais conhecidos:

Classificação Decimal de Dewey (CDD), e a Classificação Decimal Universal (CDU), a Classificação da Biblioteca do Congresso Americano (LC Library of Congress); a Classificação de Assunto de Brown (Subject Classification), a Classificação Bibliográfica de Bliss (Bibliographic Classification), e a Classificação de Dois Pontos de Ranganathan (Colon Classification).

As classificações que têm como estrutura a decimalidade mostram-se vantajosas no progresso da representação do conhecimento. A CDD é o sistema de classificação mais empregado no mundo atualmente. Segundo Felipe (2016, p. 53), “isso se deve ao fato de que na CDD é capaz de reunir, em uma mesma notação, aspectos ligados à representação temática e descritiva da obra. Por meio da CDD, se consegue dividir os assuntos de forma decimal”.

Segundo Miranda e Silva (2019, p. 534) “a CDD é um esquema de classificação bibliográfica que apresenta um sistema enumerativo, cuja classificação envolve conceitos de hierarquização de conceitos segundo disciplinas, representando a tradição científica de seu tempo”. Desenvolvida por Dewey em 1872, influenciada por elementos da filosofia de Aristóteles, Francis Bacon, Hegel e Harris, divide o conhecimento em dez disciplinas ou classes principais. Segundo Miranda e Silva (2019), a CDD se tornou um dos instrumentos de organização do conhecimento mais utilizado no mundo pela sistematização e hierarquização do conhecimento.

No Brasil, temos a Classificação Decimal de Direito (CDDir), que foi idealizada pela bibliotecária Dóris de Queiroz Carvalho na década de 40 do século passado, com o objetivo de organizar o acervo de Direito da Biblioteca do Ministério da Fazenda (BMF) (AROUCK, JAEGGER E PINHA, 2017). Dóris percebeu que a CDD, no que concerne à área do Direito, mais especificamente a classe 340, não atendia as necessidades da biblioteca e deveria ser esmiuçada, e assim ela o fez. A CDDir, assim como a CDD, tem como base a decimalidade e está organizada de forma hierárquica a partir de 4 grandes grupos: 341 - Direito Público; 342 - Direito Privado; 343 - Direito Canônico e 344 - Direito Romano. Atualmente, a CDDir é o sistema mais utilizado na área do Direito no Brasil.

Ainda em se tratando de classificações que contêm a decimalidade em sua estrutura, tem-se a Classificação Decimal Universal (CDU), elaborada pelos belgas Paul Otlet e Henri la Foutaine. A CDU baseia-se na CDD porém, no momento da classificação, são utilizados símbolos para sinalizar alguns pontos especiais de um assunto e/ou suas relações. A CDU vem se atualizando ao longo dos anos e no Brasil o responsável pela sua edição e tradução é o Instituto Brasileiro de Informação em Ciência e Tecnologia.

#### **2.4 Vocabulários Controlados**

Por sua vez, os vocabulários controlados surgem para padronizar os termos já utilizados na indexação nos sistemas de recuperação da informação, eliminando assim a ambiguidade dos termos empregados no sistema. Para a ANSI/NISO Z39.19:2005, um vocabulário controlado é:

Uma lista de termos que foram enumerados explicitamente. Esta lista é controlada e está disponível por uma autoridade de registro de vocabulário controlada. Todos os termos em um vocabulário controlado devem ter uma definição inequívoca e não redundante. (NATIONAL INFORMATION STANDARDS ORGANIZATION, 2005, p. 5, tradução nossa).

Isso quer dizer que, cada termo deve apresentar uma definição e significado para controlar a ambiguidade e sinonímia do termo. Ambiguidade ocorre quando um termo possui mais de um significado, e sinonímia quando um termo pode apresentar vários outros como semelhante para exemplificar seu significado. Conforme Harpring (2016, p.37), “as funções mais importantes de um vocabulário controlado são o agrupamento de termos variantes e sinônimos em conceitos e a ligação dos conceitos em uma ordem lógica ou sua classificação em categorias”. Evitar a ambiguidade e a sinonímia trará mais precisão no momento da busca do termo por parte do usuário. Dessa forma, os vocabulários apresentam regras a serem seguidas na sua estruturação.

Todo vocabulário controlado é composto por um conjunto de termos que representam conceitos de um ou vários campos de conhecimento. Tais signos são dispostos em estrutura relacional previamente definida. Em geral, os vocabulários controlados são apresentados em ordem hierárquica e alfabética (macroestrutura e microestrutura). (KOBASHI, 2008, p.1)



As macroestruturas são relativas às áreas do conhecimento do vocabulário, podendo ser temáticos ou não, e nelas são apresentados os conceitos gerais que compõem o vocabulário de forma mais ampla e hierárquica. Por sua vez, as microestruturas são os descritores, os termos propriamente ditos.

## 2.5 Ontologias e Taxonomias

O termo Ontologia é conhecido por sua origem na filosofia e seu uso em comum em outras áreas do conhecimento, como a Ciência da Computação, além da Ciência da Informação. Etimologicamente vem do grego e significa estudo do ser, Ontos (ser) + logos (estudo) (CURRÁS, 2010). No contexto da Ciência da Informação, uma Ontologia “pode ser vista como um vocabulário de representação, geralmente especializado em algum domínio ou assunto, qualificado por conceituações de tipos de objetos e suas relações no mundo” (FERNEDA; DIAS, 2017, p.174). Logo, uma Ontologia tem a característica de agrupar os conceitos de um determinado domínio tal como um Tesouro, porém numa Ontologia é possível observar as definições de um determinado conceito. Para Moreiro Gonzáles (2011, p. 77)

uma ontologia é uma descrição explícita e formal de conceitos em um domínio de discurso (classes, também chamadas conceitos), propriedades de cada conceito, descrevendo várias características e atributos do conceito (slots – funções ou propriedades), e restrições sobre slots (facetar – restrições de uma função).

Pode-se dizer que, uma ontologia é uma reunião de conceitos que juntos formam a representação de uma determinada área do conhecimento. As ontologias são construídas por meio de axiomas e relações. Os Axiomas são como os conceitos são estabelecidos, suas definições e as relações como esses conceitos se relacionam entre si.

Por sua vez, a Taxonomia tem sua origem na Biologia e com o advento das tecnologias passou a ser usada no cenário da informação. O termo taxonomia deriva do grego e conforme Currás (2010, p.58) “taxis = ordenação e nomia = lei, norma, regra”. Assim, tanto na Biologia quanto na informação significa classificar, arranjar.

Segundo Smiraglia (2012, p. 51, tradução nossa), “uma taxonomia é uma lista ordenada de termos junto com suas definições ou outras características determinantes. Taxonomia é uma forma de definir entidades e componentes em um domínio”. As Taxonomias são formadas de maneira hierárquica, mas não incluem as definições dos termos, como ocorre em uma Ontologia. As Taxonomias e as Ontologias podem ser aplicadas facilmente no contexto da web semântica, não sendo aplicadas em contextos em que os acervos são impressos.

Após compreender como os SOC funcionam, se faz necessário entender como as unidades de informação estão utilizando os SOC em seus processos de organização do conhecimento. O contexto escolhido para a presente pesquisa foi o Sistema de Bibliotecas da UFRJ.

### **3 As bibliotecas que compõem o SiBI da UFRJ**

Para a identificação das bibliotecas que compõem o Sistema de Bibliotecas e Informação (SiBI) da Universidade Federal do Rio de Janeiro (UFRJ) foi necessário consultar o *website* a partir da *url* <https://www.sibi.ufrj.br/index.php/bibliotecas/todas-as-bibliotecas>, na seção “Todas as bibliotecas”.

Foram identificadas 44 bibliotecas, divididas em oito Centros, a saber:

- 1) Centro de Ciências Jurídicas e Econômicas (CCJE)
- 2) Centro de Ciências Matemáticas e da Natureza (CCMN)
- 3) Centro de Ciências da Saúde (CCS)
- 4) Centro de Filosofia e Ciências Humanas (CFCH)
- 5) Centro de Letras e Artes (CLA)
- 6) Centro de Tecnologia (CT)
- 7) Fórum de Ciência e Cultura (FCC)
- 8) Macaé Campus Professor Aloísio Teixeira

O Centro de Ciências Jurídicas e Econômicas (CCJE) integra quatro bibliotecas: A Biblioteca Eugênio Gudín (BEG) do Centro de Ciências Jurídicas e Econômicas (CCJE) que atende aos cursos de graduação e pós-graduação do Instituto de Economia (IE), da Faculdade de Administração e Ciências Contábeis (FACC), do Instituto de Relações Internacionais e Defesa (IRID). A Biblioteca Prof. Agrícola Bethlem do Instituto de Pós-Graduação e Pesquisa em Administração (COPPEAD) que atende ao Programa de Mestrado em Administração do Centro de Ciências Jurídicas e Econômicas e ao Instituto de Pós-graduação e Pesquisa em Administração (COPPEAD). A Biblioteca Carvalho de Mendonça (BCM) atende a Faculdade de Direito (FND). A Biblioteca do Instituto de Pesquisa e Planejamento Urbano e Regional (IPPUR) atende ao Instituto de Pesquisa e Planejamento Urbano e Regional (IPPUR).

O Centro de Ciências Matemáticas e da Natureza (CCMN) compreende oito bibliotecas: A Biblioteca Central do Centro de Ciências Matemáticas e da Natureza (CCMN) e sete bibliotecas setoriais: a Biblioteca Plínio Sussekind Rocha do Instituto de Física (IF); a Biblioteca Professor Leopoldo Nachbin do Instituto de Matemática (IM); a Biblioteca Professor Jorge de Abreu Coutinho do Instituto de Química (IQ); a Biblioteca do Núcleo de Computação Eletrônica (NCE); a Biblioteca Professor

Sílio Vaz do Observatório do Valongo (OV); a Biblioteca Professor Maurício de Almeida Abreu do Programa de Pós-graduação em Geografia (PPGG); e o Posto de Serviço de Informação do Polo de Xistoquímica (Biblioteca de Xistoquímica).

O Centro de Ciências da Saúde (CCS) contém 14 bibliotecas: a Biblioteca Central do Centro de Ciências da Saúde (CCS); a Biblioteca da Escola de Enfermagem Anna Nery (EEAN); a Biblioteca da Faculdade de Farmácia (FF); a Biblioteca do Hospital Universitário Clementino Fraga Filho (HU); a Biblioteca do Instituto de Estudos em Saúde Coletiva (IESC); a Biblioteca do Instituto de Ginecologia (IG); a Biblioteca do Instituto de Microbiologia Professor Paulo de Góes (IMPG); a Biblioteca do Instituto de Neurologia Deolindo Couto (INDC); a Biblioteca Asdrubal Costa do Instituto de Puericultura e Pediatria Martagão Gesteira (IPPMG); a Biblioteca João Ferreira da Silva Filho do Instituto de Psiquiatria (IPUB); a Biblioteca Jorge de Rezende da Maternidade Escola (ME); a Biblioteca Aloísio Teixeira do Núcleo em Ecologia e Desenvolvimento Socioambiental de Macaé (NUPEM); a Biblioteca de Recursos Instrucionais do Núcleo de Tecnologia Educacional para Saúde (NUTES); e a Biblioteca do Campus Duque de Caxias (CDC).

O Centro de Filosofia e Ciências Humanas (CFCH) encerra três bibliotecas: a Biblioteca Central do Centro de Filosofia e Ciências Humanas (CFCH); a Biblioteca do Colégio de Aplicação da UFRJ (CAP); e a Biblioteca Marina São Paulo de Vasconcellos do Instituto de Filosofia e Ciências Sociais (IFCS).

O Centro de Letras e Artes (CLA) conta com cinco bibliotecas: a Biblioteca Professor Alfredo Galvão da Escola de Belas Artes (EBA); a Biblioteca de Obras Raras da Escola de Belas Artes (EBAOR); a Biblioteca Alberto Nepomuceno da Escola de Música (BAN); a Biblioteca Lúcio Costa da Faculdade de Arquitetura (FAU); e a Biblioteca José de Alencar da Faculdade de Letras (FL).

O Centro de Tecnologia (CT) inclui cinco bibliotecas: a Biblioteca Central do Centro de Tecnologia (CT); a Biblioteca de Obras Raras ou Antigas do Centro de Tecnologia; a Biblioteca Dirceu de Alencar Velloso (BPDV) que atende ao Programa de Engenharia Civil da COPPE; a Biblioteca Paulo Geyer da Escola de Química (EQ); a Biblioteca Professor Carlos Alberto Hemais do Instituto de Macromoléculas (IMA).

O Fórum de Ciência e Cultura (FCC) incorpora quatro bibliotecas: a Biblioteca Pedro Calmon do Fórum de Ciência e Cultura (BPC); a Biblioteca do Museu Nacional (MN); a Biblioteca Francisca Keller do Programa de Pós-Graduação em Antropologia Social (PPGAS); e a Central de Memória Acadêmica (CMA) que é uma biblioteca depositária de todas as teses e dissertações defendidas na UFRJ.

Em Macaé está a Biblioteca do Campus UFRJ Macaé Professor Aloísio Teixeira atendendo aos cursos de graduação e pós-graduação do Campus UFRJ-Macaé Professor Aloísio Teixeira e da Faculdade Municipal de Macaé Professor Miguel

Ângelo da Silva Santos (FeMASS).

Conforme descrito anteriormente, foi aplicado um questionário, no período de 11 de janeiro a 19 de fevereiro de 2021, como forma de coleta de dados. Responderam ao questionário 32 bibliotecas representativas dos oito Centros, equivalente a 72,73% do total, sendo distribuídas conforme Quadro 1.

**Quadro 1: Bibliotecas respondentes do SIBI/UFRJ separadas por Centros**

Centro	Biblioteca
CCMN	Biblioteca Professor Leopoldo Nachbin do Instituto de Matemática - IM
	Biblioteca Professor Sílilo Vaz do Observatório do Valongo - OV
	Biblioteca do Núcleo de Computação Eletrônica - NCE
CCS	Biblioteca do Instituto de Neurologia Deolindo Couto - INDC
	Biblioteca do Instituto de Microbiologia Professor Paulo de Góes - IMPG
	Biblioteca Jorge de Rezende da Maternidade Escola - ME
	Biblioteca do Hospital Universitário Clementino Fraga Filho e do Instituto de Doenças do Tórax
	Biblioteca do Instituto de Estudos em Saúde Coletiva - IESC
	Biblioteca Central do Centro de Ciências da Saúde - CCS
	Biblioteca do Instituto de Ginecologia - IG
	Biblioteca da Faculdade de Farmácia - FF
	Biblioteca Asdrubal Costa do Instituto de Puericultura e Pediatria Martagão Gesteira - IPPMG
	Biblioteca de Recursos Instrucionais do Núcleo de Tecnologia Educacional para Saúde - NUTES
CFCH	Biblioteca Marina São Paulo de Vasconcellos do Instituto de Filosofia e Ciências Sociais - IFCS
	Biblioteca Central do Centro de Filosofia e Ciências Humanas - CFCH
	Biblioteca do Colégio de Aplicação da UFRJ - CAP
FCC	Biblioteca Francisca Keller do Programa de Pós-Graduação em Antropologia Social - PPGAS
	Biblioteca do Museu Nacional - MN
	Central de Memória Acadêmica - CMA
	Biblioteca Pedro Calmon do Fórum de Ciência e Cultura - BPC
CLA	Biblioteca Alberto Nepomuceno da Escola de Música - BAN
	Biblioteca de Obras Raras da Escola de Belas Artes - EBAOR
	Biblioteca José de Alencar da Faculdade de Letras - FL
	Biblioteca Professor Alfredo Galvão da Escola de Belas Artes - EBA
CT	Biblioteca Central do Centro de Tecnologia - CT
	Biblioteca Professor Carlos Alberto Hemais do Instituto de Macromoléculas - IMA
	Biblioteca Paulo Geyer da Escola de Química - EQ
CCJE	Biblioteca Eugênio Gudín do Centro de Ciências Jurídicas e Econômicas - CCJE
	Biblioteca do Instituto de Pesquisa e Planejamento Urbano e Regional - IPPUR
	Biblioteca Carvalho de Mendonça da Faculdade Nacional de Direito - FND
	Biblioteca Prof. Agrícola Bethlem do Instituto de Pós-Graduação e Pesquisa em Administração - COPPEAD

Fonte: Elaborada pelas autoras (2021)

Comparando o total de bibliotecas por centro com o número de bibliotecas respondentes temos o seguinte percentual (Quadro 2):

**Quadro 2: Percentual de Bibliotecas Respondentes por Centro**

CENTRO	QUANTIDADE POR CENTRO	QUANTIDADE DE RESPONDENTES	PERCENTUAL
CCJE	4	4	100%
CFCH	3	3	100%
FCC	4	4	100%
MACAÉ	1	1	100%
CLA	5	4	80%
CCS	14	10	71,43%
CT	5	3	60%
CCMN	8	3	37,5%

Fonte: Elaborada pelas autoras (2021)

Dos oito Centros aos quais as bibliotecas estão vinculadas, metade foi obtida uma cobertura completa. Dos outros quatro, o Centro de Ciências Matemáticas e da Natureza (CCMN) foi o que obteve um menor percentual de respostas.

#### 4 Os Sistemas de Organização do Conhecimento das Bibliotecas do SiBI

As bibliotecas que integram o SiBI, ao catalogarem seus documentos, alimentam a Base de Dados Minerva a partir do Sistema *Aleph*<sup>5</sup>. Nela estão reunidos os catálogos *online* das 44 Bibliotecas da UFRJ a partir da *url* [www.minerva.ufrj.br](http://www.minerva.ufrj.br). Ao alimentar a base de dados, as bibliotecas também alimentam uma lista de cabeçalho de assuntos que pode ser pesquisada na aba índice que permitirá a identificação do assunto, seus desdobramentos e a quantidade de itens a ele atribuídos. As bibliotecas pesquisam na Base de Dados Minerva para controle de sinonímia e ambiguidade. Os termos ali inseridos são controlados a partir dos diferentes soc que estão sendo estudados nesta pesquisa.

As bibliotecas da UFRJ utilizam como Sistemas de Classificação Bibliográficas a Classificação Decimal de Dewey (CDD). A Biblioteca Carvalho de Mendonça utiliza, além da CDD, a Classificação Decimal de Direito (CDDir) que é específica para a área Jurídica. A Biblioteca José de Alencar desenvolveu um sistema próprio para a área de Literatura Brasileira e Portuguesa que será melhor detalhado mais adiante.

A primeira parte do questionário possui dois objetivos. O primeiro, identificar as bibliotecas respondentes, seus centros e as áreas do conhecimento atendidas e o segundo, identificar se as bibliotecas que compõem o SiBI possuem um soc construído, adaptado ou apropriado por elas. Conforme pode ser visualizado no Quadro 3.

<sup>5</sup> *Software* de gerenciamento de bibliotecas e centros de documentação

**Quadro 3: Biblioteca Respondente por Centro, área do conhecimento atendida e SOC utilizados**

Centro	Biblioteca	Áreas do conhecimento atendida	Instrumento	
CCMN	Biblioteca Professor Leopoldo Nachbin do Instituto de Matemática - IM	Matemática	Catálogo de autoridade da <i>Library of Congress</i>	
		Ciência da computação		
		História da matemática		
		Estatística		
		Atuária		
	Biblioteca Professor Sílio Vaz do Observatório do Valongo - OV	Astronomia	Catálogo de autoridade da Biblioteca Nacional Catálogo de autoridade da <i>Library of Congress</i>	
Biblioteca do Núcleo de Computação Eletrônica - NCE	Ciência da Computação	Lista de Cabeçalho de Assuntos da Base de Dados Minerva		
		Catálogo de autoridade da Biblioteca Nacional		
CCS	Biblioteca do Instituto de Neurologia Deolindo Couto - INDC	Neurologia	Descritores em Ciências da Saúde (DeCS)	
		Neurocirurgia		
		Fonoaudiologia		
		Fisioterapia		
		Neuropsicologia		
	Biblioteca do Instituto de Microbiologia Professor Paulo de Góes - IMPG	Ciências da Saúde		Ciências Biológicas
				Obstetrícia
	Biblioteca Jorge de Rezende da Maternidade Escola - ME	Perinatologia		
	Biblioteca do Hospital Universitário Clementino Fraga Filho e do Instituto de Doenças do Tórax	Área da saúde		
	Biblioteca do Instituto de Estudos em Saúde Coletiva - IESC	Saúde Pública		
	Biblioteca Central do Centro de Ciências da Saúde - CCS	Área da saúde		
	Biblioteca do Instituto de Ginecologia - IG	Medicina		Ginecologia
				Mastologia
Biblioteca da Faculdade de Farmácia - FF	Área farmacêutica			
Biblioteca Asdrubal Costa do Instituto de Puericultura e Pediatria Martagão Gesteira - IPPMG	Pediatria	Puericultura		
Biblioteca de Recursos Instrucionais do Núcleo de Tecnologia Educacional para Saúde - NUTES	Educação	Catálogo de autoridade da Biblioteca Nacional		
		Saúde		
		Educação em ciências		

**Quadro 3: Biblioteca Respondente por Centro, área do conhecimento atendida e SOC utilizados**

Centro	Biblioteca	Áreas do conhecimento atendida	Instrumento
CFCH	Biblioteca Marina São Paulo de Vasconcellos do Instituto de Filosofia e Ciências Sociais - IFCS	Ciências Sociais	Catálogo de autoridade da Biblioteca Nacional
		História	Catálogo de autoridade da <i>Library of Congress</i>
		Filosofia	Tesouro de Folclore e Cultura Popular
	Biblioteca Central do Centro de Filosofia e Ciências Humanas - CFCH	Comunicação	Lista de Cabeçalho de Assuntos do CFCH
		Educação	
		Psicologia	
		Serviço Social	
	Biblioteca do Colégio de Aplicação da UFRJ - CAP	Políticas Públicas	Lista de Cabeçalho de Assuntos da Base de Dados Minerva
		Educação	
		Literatura	
		Filosofia	
		Sociologia	
		Artes Cênicas	
		Cinema	
		Educação artística	
		Ciências	
		Matemática	
FCC	Biblioteca Francisca Keller do Programa de Pós-Graduação em Antropologia Social - PPGAS	Ciências Sociais	Catálogo de autoridade da Biblioteca Nacional
		Antropologia Social	Lista de Cabeçalho de Assuntos da Base de Dados Minerva
	Biblioteca do Museu Nacional - MN	Ciências Naturais	Catálogo de autoridade da <i>Library of Congress</i>
		Antropológicas	
	Central de Memória Acadêmica - CMA	Todas as áreas do conhecimento	Lista de Cabeçalho de Assuntos da Base de Dados Minerva
	Biblioteca Pedro Calmon do Fórum de Ciência e Cultura - BPC	História da Cidade do Rio de Janeiro	Catálogo de autoridade da Biblioteca Nacional
		História do Brasil	
Ciências Sociais			
		Ciências Humanas	

**Quadro 3: Biblioteca Respondente por Centro, área do conhecimento atendida e SOC utilizados**

Centro	Biblioteca	Áreas do conhecimento atendida	Instrumento
CLA	Biblioteca Alberto Nepomuceno da Escola de Música - BAN	Música	Catálogo de autoridade da Biblioteca Nacional
	Biblioteca de Obras Raras da Escola de Belas Artes - EBAOR	Artes	Catálogo de autoridade da Biblioteca Nacional
		Arquitetura	Catálogo de autoridade da <i>Library of Congress</i>
	Biblioteca José de Alencar da Faculdade de Letras - FL	Todas as áreas do conhecimento	Tabela de Classificação própria para área de Literatura
			Catálogo de autoridade da Biblioteca Nacional
			Catálogo de autoridade da <i>Library of Congress</i>
			Catálogo das Bibliotecas Universitárias Brasileiras
Biblioteca Professor Alfredo Galvão da Escola de Belas Artes - EBA	Artes	Catálogo de autoridade da Biblioteca Nacional	
		Catálogo de autoridade da <i>Library of Congress</i>	
CT	Biblioteca Central do Centro de Tecnologia - CT	Engenharia	Catálogo de autoridade da Biblioteca Nacional
		Tecnologia	Vocabulários Controlados da USP
	Catálogo de Bibliotecas Universitárias Brasileiras (UFRGS, UNICAMP, UNESP)		
	Biblioteca Professor Carlos Alberto Hemais do Instituto de Macromoléculas - IMA	Química	Lista de Cabeçalho de Assuntos da Base de Dados Minerva
Biblioteca Paulo Geyer da Escola de Química - EQ	Engenharia Química	Tabela de áreas do conhecimento do CNPq	
Termos livre			

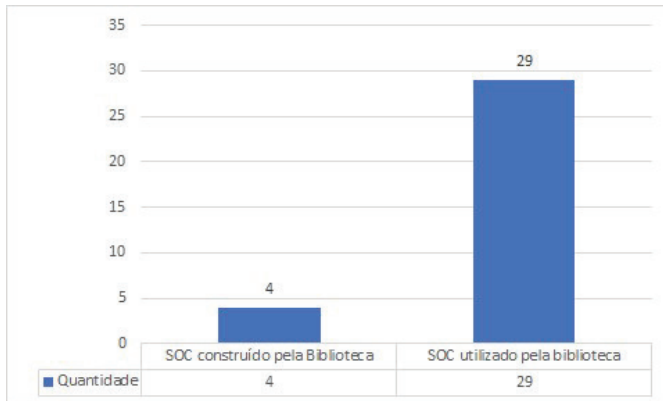


**Quadro 3: Biblioteca Respondente por Centro, área do conhecimento atendida e SOC utilizados**

Centro	Biblioteca	Áreas do conhecimento atendida	Instrumento
CCJE	Biblioteca Eugênio Gudin do Centro de Ciências Jurídicas e Econômicas - CCJE	Economia	Lista de Cabeçalho de Assuntos das Bibliotecas do CCJE
			Lista de Cabeçalho de Assuntos da Base de Dados Minerva
		Administração	Catálogo das Bibliotecas Universitárias Brasileiras
			Catálogo de autoridade da Biblioteca Nacional
		Biblioteconomia	Catálogo de autoridade da <i>Library of Congress</i>
			Vocabulário controlado básico do Senado
		Ciências Contábeis	Catálogo da Biblioteca da Comissão de Valores Mobiliários
			Artigos
	Relações Internacionais	Glossários	
		Dicionários	
	Defesa e Gestão Estratégica Internacional	Tesouros	
	Biblioteca do Instituto de Pesquisa e Planejamento Urbano e Regional - IPPUR	Planejamento Urbano e Regional	Lista de Cabeçalho de assuntos criada pelo IPPUR
	Biblioteca Carvalho de Mendonça da Faculdade Nacional de Direito - FND	Direito	Catálogos de assunto das Bibliotecas Nacionais
			Catálogo de autoridade da <i>Library of Congress</i>
			Vocabulário controlado básico do Senado
Biblioteca Prof. Agrícola Bethlem do Instituto de Pós-Graduação e Pesquisa em Administração - COPPEAD	Administração	Lista de Cabeçalho de Assuntos da Base de Dados Minerva	
		Lista de Cabeçalhos de Assunto do CCJE	

Fonte: Elaborada pelas autoras (2021)

O questionário trouxe como resultado diferentes tipos de SOC utilizados pelas bibliotecas da UFRJ. Das 32 bibliotecas respondentes, 12,5% construíram seus SOC e 90,62% utilizam algum SOC para controle de vocabulário, conforme gráfico 1. Vale ressaltar que a Biblioteca José de Alencar da Faculdade de Letras construiu um SOC de sistema de classificação e utiliza SOC para controle de vocabulário, desta forma ela está inserida nos dois percentuais.

**Gráfico 1: SOC construídos e utilizados pelas bibliotecas que compõem o SiBI**

Fonte: Elaborado pelas autoras (2021)

Dos quatro soc construídos pelas bibliotecas, dois estão em formato analógico. Todos os utilizados são em formato digital.

Em termos quantitativos, os soc mais utilizados pelas bibliotecas que compõem o SiBI são o Catálogo de autoridade da Biblioteca Nacional, citado 12 vezes, seguido pelo DeCS citado 10 vezes e pelo Catálogo de autoridade da *Library of Congress* citado 9 vezes, conforme quadro 4.

**Quadro 4: Quantitativo de SOC utilizados pelas Bibliotecas que compõem o SiBI/UFRJ**

SOC	Quantidade
Catálogo de autoridade da Biblioteca Nacional	12
Descritores em Ciências da Saúde (DeCS)	10
Catálogo de autoridade da <i>Library of Congress</i>	9
Catálogo das Biblioteca Universitárias Brasileiras	2
Vocabulários Controlados da USP	2
Vocabulário Controlado Básico do Senado Federal	2
Tesouro de Folclore e Cultura Popular	1
Thesaurus Brasileiro da Educação do INEP	1
Catálogo das Bibliotecas Nacionais do Mundo	1
Catálogo da Biblioteca da UFRGS	1
Catálogo da Biblioteca da UNICAMP	1
Tesouro da UNESP.	1
Tabela de áreas do conhecimento do CNPq	1
Catálogo on-line da Biblioteca da Comissão de Valores Mobiliários	1
Catálogo das Bibliotecas Nacionais do Mundo	1

Fonte: Elaborado pelas autoras (2021)

Para análise detalhada das características dos SOC, eles foram divididos seguindo a tipologia de Zeng (2008), em que são organizados a partir da complexidade das estruturas e principais funções.

#### 4.1 Listas de Cabeçalhos de Assunto

**Lista de Cabeçalho de assuntos do CCJE** é uma lista de descritores que teve sua construção inicializada, em 1980, com o intuito de normalizar e unificar os cabeçalhos de assunto das quatro bibliotecas que compõem o Centro de Ciências Jurídicas e Econômicas (CCJE). A coleta dos termos utilizou a lista individual das Bibliotecas do CCJE, COPPEAD, IPPUR e a FND entrou posteriormente, em 1994. A união das listas gerou uma lista padronizada, em dois volumes (A-I, J-Z). Além desta listagem foram criadas: a Lista de Subcabeçalhos e Identificadores; de Entidades coletivas; e de Identificadores geográficos. Para a padronização desses termos, o grupo criou uma política de indexação unificada. A pesquisa dos termos era realizada em fontes apropriadas para cada área do conhecimento, sendo registrado no controle de vocabulário qual foi a biblioteca responsável pela entrada do termo novo (autoridade), qual foi a fonte pesquisada (Fonte consultada), as fontes básicas e as fontes dispensadas. Os assuntos eram subdivididos por forma, local e tempo e, quando necessário, era incluído um termo qualificador. Os termos são apresentados em português e inglês. Em 2019 a lista foi transcrita para uma planilha de *Excel*, que deverá ser apresentada nas bibliotecas que compõem o centro do CCJE para atualização, manutenção e validação do SOC.

**Lista de Cabeçalho de assuntos do CFCH** é uma lista de descritores criada pela equipe da biblioteca entre os anos de 1987 e 1989. Sua utilização pela equipe de processamento técnico teve início em 1994. A ordem de apresentação dos termos é a alfabética. Os termos são extraídos a partir da análise conceitual das teses, dissertações, artigos de periódicos ou livros. Termos novos são inseridos em um formulário em que a equipe de catalogação analisa, conceitua e os valida para inserção na listagem. Os conceitos são estabelecidos com base no critério de pesquisa em tesouros especializados, dicionários, enciclopédias, CDD, base de dados especializadas e quando os termos são muito novos são consultados os professores, técnicos da área, pesquisadores da área e fontes da internet. Inicialmente estava apenas disponibilizada em formato analógico sendo hoje disponível em *Excel* por um link que apenas a equipe tem acesso. Foi criada com o propósito de “estabelecer um melhor controle e padronização dos termos utilizados no momento da indexação” (BIBLIOTECA DO CFCH, 2021).

**Lista de Cabeçalho de assuntos do IPPUR** é uma lista de descritores criada por bibliotecários há mais de dez anos e que sofre constante atualização. Os conceitos são inseridos conforme os termos surgem nos trabalhos acadêmicos dos alunos da unidade, estes são pesquisados no catálogo de autoridade da Biblioteca Nacional e os termos autorizados são acrescentados à lista. Este instrumento está disponibilizado em formato analógico e no idioma português. O propósito de criação da listagem foi para “atualizar os termos para que se adequem aos assuntos utilizados na área, com base na Biblioteca Nacional” (BIBLIOTECA DO IPPUR, 2021).

**Catálogo de autoridade da Biblioteca Nacional (BN)** reúne descritores para nomes de pessoas, entidades coletivas, eventos associados à autoria de obras e assuntos. A recuperação das pesquisas dos termos autorizados apresenta: termos gerais (TG), termos específicos (TE) e termos relacionados (TR). Engloba tópicos, remissivas ver, remissivas ver também, além das subdivisões gerais, cronológicas e geográficas. Traz também a fonte positiva dos dados e a nota geral pública.

Link: [http://acervo.bn.gov.br/sophia\\_web/busca/autoridades](http://acervo.bn.gov.br/sophia_web/busca/autoridades)

**Catálogo de autoridade da *Library of Congress* (LC)** reúne descritores autorizados para assunto, nomes pessoais, títulos, nomes e títulos e palavras-chave. O resultado da busca apresenta quatro colunas que trazem se o título é autorizado, quantos registros bibliográficos têm no catálogo da biblioteca com esse assunto, o cabeçalho de assunto e o vocabulário controlado utilizado para padronização. Ao clicar no ícone do termo desejado abre-se uma nova tela com: título, referências e/ou notas de escopo indicadas pelo ícone selecionado.

Link: <http://authorities.loc.gov/>

**Descritores em Ciências da Saúde (DeCs)** surgiu em 1970 a partir da autorização do *National Library of Medicine* (NLM) dos Estados Unidos para que a Bireme traduzisse e adaptasse o vocabulário do *Medical Subject Headings* (MeSh) para o português e espanhol. Sua estrutura é hierárquica, com atualização anual. Além dos termos usados no MeSh foram incluídas mais quatro categorias: Saúde Pública, Homeopatia, Ciência e Saúde e Vigilância Sanitária. Faz parte do Sistema Unificado de Linguagem Médica (UMLS) buscando a interoperabilidade entre os sistemas da área de Saúde e Biomedicina. O resultado de busca do termo exibe três tipos de apresentação: **Detalhes** em que são relacionados os descritores nos quatro idiomas, termo alternativo, nota de escopo, nota de indexação, qualificadores permitidos; **Estrutura hierárquica** em que o conceito aparece na árvore conceitual do DeCS/MeSH; **Conceitos** em que são apresentados o conceito preferido e o conceito mais

específico podem trazer ou não a nota de escopo e o termo alternativo.

Link: <https://decs.bvsalud.org/>

**Catálogo da Biblioteca da Comissão de Valores Mobiliários (CVM)** reúne descritores para nome de pessoas, assuntos e editoras. A recuperação das pesquisas dos termos autorizados apresenta: remissiva ver (US), termo geral (TG) e termo relacionado (TR).

Link: [http://biblioteca.cvm.gov.br/scripts/bnportal/bnportal.exe/index#a-cao=lista&view=vbiblasso&order=tit1&sv=vbiblasso1:remissivas:tipo\\_rel%20desc,class\\_ord&alias=geral](http://biblioteca.cvm.gov.br/scripts/bnportal/bnportal.exe/index#a-cao=lista&view=vbiblasso&order=tit1&sv=vbiblasso1:remissivas:tipo_rel%20desc,class_ord&alias=geral)

**Catálogo das Bibliotecas Universitárias Brasileiras** utilizado como fonte de pesquisa para identificação das terminologias utilizadas no meio acadêmico de determinada área. Sendo citadas pelas bibliotecas respondentes, as seguintes:

**Lista de Cabeçalho de Assuntos da base SABI - UFRGS** reúne descritores utilizados pelas bibliotecas que compõem o Sistema de Bibliotecas da Universidade Federal do Rio Grande do Sul (UFRGS). A recuperação dos termos apresenta uma relação dos termos autorizados e a quantidade de itens atribuídos a esses termos.

Link: <https://sabi.ufrgs.br>

**Catálogo de autoridade da Base Acervus - UNICAMP** reúne descritores para nomes de pessoas, entidades coletivas, eventos associados à autoria de obras e assuntos. A recuperação das pesquisas dos termos autorizados apresenta: remissiva ver (US/UF), remissiva ver também (TR), fonte positiva dos dados, nota geral pública e o termo no idioma inglês.

Link: <http://acervus.unicamp.br/>

**Catálogo das Bibliotecas Nacionais do Mundo** utilizado pelas bibliotecas para consulta de autoridades (autoria) e de assunto. Não foi citado exemplos pelas bibliotecas respondentes.

#### 4.2 Esquemas de Classificação Bibliográfica

**Tabela de Classificação para área de Literatura Brasileira e Portuguesa** é uma tabela adaptada da CDD destinada a classificação dos livros de literatura brasileira. Criada por bibliotecários há mais de 40 anos. Os novos termos são incluídos na tabela na “tabela de classificação para números próprios (autor) e/ou novos assuntos” (BIBLIOTECA DE LETRAS, 2021). O idioma da tabela é o português e ela está

disponível em formato analógico. Esta tabela foi construída com o propósito de suprir as necessidades deixadas pela tabela da CDD de classificar “os livros da área de literatura brasileira e portuguesa” (BIBLIOTECA DE LETRAS, 2021).

**Tabela de áreas do conhecimento CNPq (Tabela CNPq)** é uma classificação desenvolvida pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) com a finalidade de sistematizar informações sobre o desenvolvimento científico e tecnológico. A tabela é dividida em quatro níveis: grande área, área do conhecimento, subárea e especialidade.

- Link: [https://www.gov.br/capes/pt-br/centrais-de-conteudo/TabelaAreas-Conhecimento\\_072012\\_atualizada\\_2017\\_v2.pdf](https://www.gov.br/capes/pt-br/centrais-de-conteudo/TabelaAreas-Conhecimento_072012_atualizada_2017_v2.pdf)

#### 4.3 Vocabulários Controlados

**Vocabulário controlado da USP (vc USP)** é uma lista de assuntos criada para dar suporte ao Banco de Dados Bibliográficos da Universidade de São Paulo – DEDALUS. O vocabulário é multidisciplinar, abrangendo as diferentes áreas atendidas pelas bibliotecas da USP. O vocabulário pode ser consultado pela Macroestrutura; pela lista de assuntos dividida em: geral alfabética, geral hierárquica, locais geográficos e históricos, gênero e forma, profissões e ocupações; e pela tabela de qualificadores.

- Link: <http://vocabusp.sibi.usp.br/vocab/>

**Vocabulário controlado básico do Senado (vcB Senado Federal)** é um vocabulário controlado com relações hierárquicas, associativas, partitivas e/ou de equivalências terminológicas. Foi criado na década de 1980 para manter a uniformidade da indexação e da recuperação da informação dos documentos bibliográficos da Rede Virtual de Bibliotecas (RVBI). Sua cobertura inclui diversas áreas do conhecimento, priorizando as Ciências Sociais e Humanas com ênfase no Direito. Sua atualização é constante. Os resultados da pesquisa do termo apresentam assunto, usado para, termo geral, termo específico, termo relacionado, definição, equivalência, categoria e notação para CDD e CDDir.

- Link: [http://biblioteca2.senado.gov.br:8991/F/?func=find-b-o&local\\_base=senio](http://biblioteca2.senado.gov.br:8991/F/?func=find-b-o&local_base=senio)

#### 4.4 Tesouros

**Thesouro Brasileiro de Educação (BRASED)** é um vocabulário controlado que agrupa termos e conceitos. Esses termos são chamados de descritores e se relacionam entre si a partir de sua estrutura conceitual, sendo extraídos de documentos analisados pelo Centro de Informação e Biblioteca em Educação (Cibec). Os ter-

mos pesquisados apresentam como resultado: uma estrutura das relações hierárquicas e a estrutura do termo. A estrutura do termo traz: Termo, TA, Conceituações, Documentos indexados.

- Link: <http://inep.gov.br/thesaurus-brasileiro-da-educacao>

**Tesouro de Folclore e Cultura Popular Brasileiro** desenvolvido pelo Centro Nacional de Folclore e Cultura Popular do Instituto do Patrimônio Histórico e Artístico Nacional/MinC (IPHAN), com patrocínio do Programa Caixa de Adoção de Entidades Culturais, da Caixa Econômica Federal. A equipe que fez parte do seu desenvolvimento é multidisciplinar com vasta experiência na área da cultura popular brasileira. Os termos foram levantados a partir dos acervos da Biblioteca Amadeu Amaral e do Museu de Folclore Edison Carneiro. Os termos apresentam relações entre eles (relação de equivalência, relação hierárquica, relação partitiva e relação associativa) e notas explicativas (notas de aplicação e notas de indexação). A apresentação dos termos está estruturada de duas formas: sistemática (ordem classificatória) e alfabética.

- Link: <http://www.cnfcp.gov.br/tesouro/>

**Tesouro UNESP** reúne descritores organizados semanticamente e genericamente. Sua atualização ocorre desde 2013 sob a responsabilidade do Grupo de Linguagem UNESP. Este grupo é regulamentado pela Portaria CGB/UNESP nº 01 de 16 de maio de 2017 e tem como integrantes catalogadores das bibliotecas e pesquisadores da área de Organização do Conhecimento. Os resultados da pesquisa do termo apresentam termos não preferidos (UP), términos genéricos (TG), términos específicos (TE), termos relacionados (TR), nota de escopo e os metadados que incluem a data de criação do termo e de inclusão no tesouro.

- Link: <https://www.biblioteca.unesp.br/tesouro/vocab/index.php>

#### 4.5 Síntese sobre os SOC em uso no SiBI da UFRJ

Na presente seção apresentaremos as características relativas à estrutura e funcionamento dos Sistemas de Organização do Conhecimento empreendidos nas bibliotecas que compõem o SiBI da UFRJ. Será realizada uma análise das características apresentadas pelos SOC escolhidos pelas bibliotecas do SiBI com as características de cada tipologia de SOC descrita na literatura de Organização do Conhecimento.

Desta forma, no que tange aos SOC construídos pelas bibliotecas do SiBI, a análise fica limitada, pois não foi possível durante esta pesquisa ter acesso ao instrumento. Toda a descrição do mesmo foi pautada nos relatos obtidos pela resposta do

questionário. Pela descrição dos três instrumentos para controle de vocabulários, todos são uma lista de cabeçalho de assunto organizadas em ordem alfabética. São destinadas às áreas de suas bibliotecas. Somente a Lista de Cabeçalhos de assuntos do CCJE utiliza termos em português e inglês. Quanto à atualização, o Cabeçalho de assunto do IPPUR e do CFCH estão atualizados, já o do CCJE está passando por um processo de atualização e validação. A tabela de classificação criada pela Biblioteca José de Alencar da Faculdade de Letras se enquadra como um sistema de classificação e segue a estrutura da CDD com um recorte para as áreas de literatura brasileira e portuguesa. Quanto a atualização, esta ocorre constantemente.

Neste contexto, os sistemas mais utilizados pelas bibliotecas são os cabeçalhos de assuntos, ao todo seis cabeçalhos apareceram nos resultados da pesquisa. Abaixo, no quadro 5, a apresentação das características dos cabeçalhos de assuntos manuseados no SiBI UFRJ.

**Quadro 5: Características dos Cabeçalho de assuntos utilizados pelas Bibliotecas da UFRJ**

CARACTERÍSTICAS	DeCs	BN	LC	CVM	SABI - UFRGS	ACERVUS - UNICAMP
ÁREA	Saúde	Multi-disciplinar	Multi-disciplinar	Mercado de Capitais e Legislação Societária	Multi-disciplinar	Multi-disciplinar
TERMO	SIM	SIM	SIM	SIM	SIM	SIM
RELAÇÃO HIERÁRQUICA (TG, TE)	SIM	SIM	SIM	SIM	NÃO	NÃO
RELAÇÃO ONTOLÓGICA (TR OU TA)	SIM	SIM	SIM	SIM	NÃO	SIM
NOTA DE ESCOPO	SIM	SIM	SIM	NÃO	NÃO	NÃO
DEFINIÇÃO	NÃO	NÃO	NÃO	NÃO	NÃO	SIM
LÍNGUAS	PORT./INGL./ESP./FRAN.	PORT./INGL.	INGL.	PORT.	PORT.	PORT./INGL.
ATUALIZAÇÃO	SIM	SIM	SIM	SIM	SIM	SIM
NOTAÇÃO	NÃO	NÃO	NÃO	NÃO	NÃO	NÃO

Fonte: Elaborado pelas autoras (2021)

Como se pode observar no quadro 5, a maioria dos cabeçalhos de assuntos são multidisciplinares, apenas o DeCs está voltado para a área da saúde e o CVM engloba os assuntos relacionados a Mercado de Capitais e Legislação Societária. Todos os cabeçalhos citados apresentam termos em sua estrutura, porém nem todos apresentam as relações hierárquicas e ontológicas.



Mesmo tendo sua origem no Brasil, dois cabeçalhos englobam termos em inglês, além do português que é a língua original das instituições responsáveis pela sua elaboração, o cabeçalho da Biblioteca Nacional e o UNICAMP. Por ser uma adaptação de um cabeçalho internacional e atender a América Latina e o Caribe, o DeCs apresenta seus termos em quatro idiomas: português, inglês, espanhol e francês.

O único que apresenta a definição junto ao termo é o cabeçalho da UNICAMP. A apresentação da definição do termo é uma escolha da instituição, não se encontra na literatura algo explanando sobre a obrigatoriedade da apresentação de conceitos nos cabeçalhos de assuntos.

Todos os cabeçalhos citados passaram por atualizações nos últimos cinco anos, como pode ser visto nas suas respectivas páginas na internet. Nenhum deles apresenta a notação ou sugestões de notação baseadas em algum sistema bibliográfico.

Acerca dos vocabulários controlados, apenas dois foram citados pelos bibliotecários: o vocabulário da USP e o vocabulário do Senado Federal. A seguir, no quadro 6, a apresentação das suas características como Sistemas de Organização do Conhecimento.

**Quadro 6: Características dos Vocabulários controlado utilizados pelas Bibliotecas da UFRJ**

CARACTERÍSTICAS	VC USP	VCB DO SENADO FEDERAL
ÁREA	Multidisciplinar	Direito e Ciências Sociais
TERMO	SIM	SIM
RELAÇÃO HIERÁRQUICA (TG, TE)	NÃO	SIM
RELAÇÃO ONTOLÓGICA (TR OU TA)	NÃO	SIM
NOTA DE ESCOPO	NÃO	NÃO
DEFINIÇÃO	NÃO	SIM
LÍNGUAS	PORT.	PORT.
ATUALIZAÇÃO	SIM	SIM
NOTAÇÃO	NÃO	SIM

Fonte: Elaborado pelas autoras (2021)

Os dois vocabulários manuseados pelas bibliotecas foram desenvolvidos no Brasil e são apresentados em português. Passaram por atualizações nos últimos cinco anos e estão disponíveis em ambiente *web*. Nenhum dos dois apresenta nota de escopo, porém o vocabulário do Senado Federal apresenta definição dos termos que o compõem, auxiliando no entendimento da área do conhecimento específica.

O vocabulário do Senado ainda apresenta a notação que deve ser aplicada junto com o termo escolhido para a indexação. Essa notação está baseada na CDD ou na CDDir, como vem demonstrado em sua apresentação. A notação junto com o

termo facilita o trabalho do bibliotecário, visto que a notação só confirma o termo escolhido.

O vocabulário da USP em sua capa aponta a existência de uma relação hierárquica entre os termos, porém ao consultarmos o vocabulário em busca de termos, os termos não apresentam essa relação e também não apresentam relação ontológica.

No que diz respeito aos tesouros, como exposto acima, foram citados apenas três tesouros como Sistemas de Organização do Conhecimento pelas bibliotecas que integram o SiBI UFRJ. Abaixo, no quadro 7, a apresentação das suas características com relação ao seu funcionamento e estrutura

**Quadro 7: Características dos Tesouros utilizados pelas Bibliotecas da UFRJ**

CARACTERÍSTICAS	TESAURO UNESP	TESAURO DE FOLCLORE E CULTURA POPULAR BRASILEIRO	BRASED
ÁREA	Multidisciplinar	Interdisciplinar	Educação
TERMO	SIM	SIM	SIM
RELAÇÃO HIERÁRQUICA (TG, TE)	SIM	SIM	SIM
RELAÇÃO ONTOLÓGICA (TR OU TA)	SIM	SIM	SIM
NOTA DE ESCOPO	SIM	SIM	NÃO
DEFINIÇÃO	NÃO	SIM	SIM
LÍNGUAS	PORT.	PORT.	PORT.
ATUALIZAÇÃO	SIM	SIM	SIM
NOTAÇÃO	NÃO	NÃO	NÃO

Fonte: Elaborado pelas autoras (2021)

Todos os tesouros citados pelas bibliotecas do SiBI estão estruturados em português e foram elaborados no Brasil. Apenas um tesouro é disciplinar, BRASED, da área da Educação. Todos os três estão disponíveis e só podem funcionar no ambiente da Web. Isso facilita também o aspecto relacionado a atualização dos termos, todos eles passaram por atualizações nos últimos cinco anos.

Nenhum deles possui notação, visto que não é uma característica dos tesouros, mas apresentam os termos e suas relações tanto hierárquica quanto ontológica, o que auxilia o bibliotecário no momento de escolher o termo para representar o conteúdo do documento. Apenas um tesouro, o BRASED, não apresenta nota de escopo, acreditamos que por ser disciplinar os organizadores não acharam que houvesse necessidade de apresentar os limites semânticos dos termos, visto que todo o conteúdo é relacionado a educação e o risco de ocorrer termos com sinonímia é baixo.

As bibliotecas citam a tabela de classificação do CNPq como um instrumento utilizado para classificar seus assuntos, apesar da mesma ter sido criada com a finalidade de gestão e avaliação da produção científica do país, este instrumento pode

auxiliar na classificação do conhecimento. Como descrito anteriormente, é uma classificação hierárquica, dividida em quatro níveis: grandes áreas, áreas, subáreas e especialidades. Não possui nota de escopo, definição ou alguma instrução de uso como pode ser encontrada na CDD. Não apresenta também relações ontológicas ou facetadas. Isto se justifica por seu propósito ser destinado a gestão e avaliação, ou seja, ela não é voltada para o controle de vocabulário ou a organização do acervo na estante.

## 5 Discussão sobre a necessidade de criação de um Sistema de Organização do Conhecimento próprio para as bibliotecas do SiBI

Conforme descrito anteriormente, as três últimas perguntas da terceira seção do questionário tiveram como objetivo identificar se as bibliotecas que não construíram um instrumento próprio para o controle do seu vocabulário, sentem a necessidade de construí-lo, por qual motivo e se pretendem construir.

Das 29 bibliotecas respondentes que não construíram seus soc, quatro não responderam a essa pergunta do questionário. Das que responderam 12 sentem a necessidade de criar um soc para sua biblioteca, 13 não sentem essa necessidade. Das que sentem a necessidade, sete estão dispostas a criar, conforme Quadro 8.

**Quadro 8: Avaliação da necessidade de criação de SOC próprios pelas Bibliotecas do SiBI**

Centro	Biblioteca	Necessita?	Pretende?
CCMN	Biblioteca Professor Leopoldo Nachbin do Instituto de Matemática - IM	SIM	NÃO
	Biblioteca Professor Sílio Vaz do Observatório do Valongo - OV	Não respondeu	Não respondeu
	Biblioteca do Núcleo de Computação Eletrônica - NCE	NÃO	NÃO
CCS	Biblioteca do Instituto de Neurologia Deolindo Couto - INDC	Não respondeu	Não respondeu
	Biblioteca do Instituto de Microbiologia Professor Paulo de Góes - IMPG	SIM	SIM
	Biblioteca Jorge de Rezende da Maternidade Escola - ME	SIM	SIM
	Biblioteca do Hospital Universitário Clementino Fraga Filho e do Instituto de Doenças do Tórax	NÃO	NÃO
	Biblioteca do Instituto de Estudos em Saúde Coletiva - IESC	NÃO	NÃO
	Biblioteca Central do Centro de Ciências da Saúde - CCS	SIM	NÃO
	Biblioteca do Instituto de Ginecologia - IG	NÃO	NÃO
	Biblioteca da Faculdade de Farmácia - FF	NÃO	NÃO
	Biblioteca Asdrubal Costa do Instituto de Puericultura e Pediatria Martagão Gesteira - IPPMG	NÃO	Não respondeu
	Biblioteca de Recursos Instrucionais do Núcleo de Tecnologia Educacional para Saúde - NUTES	NÃO	NÃO
	Biblioteca Marina São Paulo de Vasconcelos do Instituto de Filosofia e Ciências Sociais - IFCS	NÃO	NÃO
CFCH	Biblioteca do Colégio de Aplicação da UFRJ - CAP	SIM	SIM

**Quadro 8: Avaliação da necessidade de criação de SOC próprios pelas Bibliotecas do SiBI**

Centro	Biblioteca	Necessita?	Pretende?
<b>FCC</b>	Biblioteca Francisca Keller do Programa de Pós-Graduação em Antropologia Social - PPGAS	SIM	SIM
	Biblioteca do Museu Nacional - MN	NÃO	NÃO
	Central de Memória Acadêmica - CMA	Não respondeu	Não respondeu
	Biblioteca Pedro Calmon do Fórum de Ciência e Cultura - BPC	NÃO	NÃO
<b>CLA</b>	Biblioteca Alberto Nepomuceno da Escola de Música - BAN	SIM	SIM
	Biblioteca de Obras Raras da Escola de Belas Artes - EBAOR	SIM	SIM
	Biblioteca Professor Alfredo Galvão da Escola de Belas Artes - EBA	NÃO	NÃO
	Biblioteca José de Alencar da Faculdade de Letras - FL	Não respondeu	Não respondeu
<b>CT</b>	Biblioteca Central do Centro de Tecnologia - CT	SIM	NÃO
	Biblioteca Professor Carlos Alberto Hemais do Instituto de Macromoléculas - IMA	SIM	SIM
	Biblioteca Paulo Geyer da Escola de Química - EQ	SIM	Não respondeu
<b>CCJE</b>	Biblioteca Carvalho de Mendonça da Faculdade Nacional de Direito - FND	SIM	NÃO
	Biblioteca Prof. Agrícola Bethlem do Instituto de Pós-Graduação e Pesquisa em Administração - COPPEAD	NÃO	NÃO

Fonte: Elaborada pelas autoras (2021)

Apesar das diferentes áreas atendidas pelas bibliotecas, o motivo que leva as 12 bibliotecas a sentirem a necessidade de criar um soc próprio se resume ao fato dos soc utilizados não atenderem plenamente suas necessidades, sendo preciso uma melhor adequação terminológica para as áreas atendidas e a padronização dos termos.

## 6 Considerações finais

Esta pesquisa teve como objetivo investigar quais são os sistemas de organização do conhecimento utilizados pelas bibliotecas que integram o SiBI da UFRJ. Para isso foi inicialmente caracterizado os Sistemas de Organização do Conhecimento constituintes da Organização do Conhecimento.

Conforme visto, a UFRJ é uma das universidades mais antigas do país e possui em sua estrutura um Sistema de Bibliotecas e Informação (SiBI) responsável pela gestão das 44 bibliotecas que atendem aos oito Centros que integram a estrutura da universidade.

O acervo destas 44 bibliotecas está catalogado e indexado na Base Minerva que pode ser acessada pela *web*. Foi possível identificar que as 32 bibliotecas respondentes cobrem juntas as oito grandes áreas do conhecimento da Capes: Ciências Exatas e da Terra, Ciências Biológicas, Engenharias, Ciências da Saúde, Ciências Agrárias, Ciências Sociais Aplicadas, Ciências Humanas, e Linguística, Letras e Artes. Desta forma, o controle terminológico dessa indexação é de vital importância para que docentes, discentes, pesquisadores e demais públicos da universidade possam ter

acesso ao acervo e com isso recuperem a informação desejada.

Quanto a este aspecto, esta pesquisa identificou que, para controle terminológico, três dessas bibliotecas construíram seu próprio instrumento, sendo este identificado como uma lista de cabeçalho de assunto. A Biblioteca José de Alencar da Faculdade de Letras necessitou construir um instrumento para sistema de classificação, ou seja, para melhor adequar o seu acervo na estante, visto que a CDD, instrumento utilizado na universidade, não estava respondendo às suas expectativas. A Biblioteca Carvalho de Mendonça da Faculdade de Direito também utiliza a CDDir para a área jurídica, pois esta responde melhor às suas necessidades.

Outro fato identificado pela pesquisa é que os sistemas mais utilizados pelas bibliotecas respondentes são os cabeçalhos de assuntos, seguidos dos tesouros e dos vocabulários de assunto. A maioria das bibliotecas utilizam mais de um tipo de SOC, mas todas visam o controle terminológico, a eliminação das sinônimas e das ambiguidades.

### Referências

AROUCK, Osmar; JAEGGER, Fátima; PINHA, Stelina Maria. Classificação Decimal de Direito: revisão e atualização. **Revista Brasileira de Biblioteconomia e Documentação**, v. 13, n. esp. CBBDD 2017. Disponível em: <https://rbbd.febab.org.br/rbbd/article/view/954> Acesso em: 22 fev. 2021

BARITÉ ROQUETA, Mario. Sistemas de organización del conocimiento: una tipología actualizada. **Informação & Informação**, Londrina, v. 16 n. 3, p. 122 – 139, jan./ jun. 2011. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/9952>. Acesso em: 13 dez. 2020.

BRÄSCHER, Marisa; CARLAN, Eliana. Sistemas de organização do conhecimento: antigas e novas linguagens. In: ROBREDO, Jaime; BRÄSCHER, Marisa (Orgs.). **Passeios no Bosque da Informação: Estudos sobre Representação e Organização da Informação e do Conhecimento**. Brasília DF: IBICT, 2010, 335 p. ISBN: 978-85-7013-072-3. Cap. 8, p. 147-176 Edição eletrônica. Disponível em: <https://ridi.ibict.br/bitstream/123456789/36/1/eroic.pdf>. (Edição comemorativa dos 10 anos do Grupo de Pesquisa EROIC).

\_\_\_\_\_. Sistemas de Organização do Conhecimento na visão da Ciência da Informação. **RICI: Revista Ibero-americana de Ciência da Informação**, ISSN 1983-5213, Brasília, v. 4, n. 2, p. 53-73, ago./dez.2011. Disponível em: <https://periodicos.unb.br/index.php/RICI/article/view/1675>. Acesso em: 10 jan. 21.

CESARINO, Maria Augusta da Nóbrega; PINTO, Maria Cristina Mello Ferreira. Cabeçalho de assunto como linguagem de indexação. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 7, n. 2, p. 268-288, set. 1978

CLARKE, Stella G. Dextre. Thesaurus (for information retrieval). *In*: HJØRLAND, Birger; GNOLI, Claudio (Eds.) **ISKO Encyclopedia of Knowledge Organization**. Disponível em: <https://www.isko.org/cyclo/thesaurus#2>. Acesso em 10 fev. 2021.

CURRÁS, Emilia. **Ontologias, taxonomia e tesouros em teoria de sistemas e sistemática**. Tradução Jaime Robredo. Brasília: Thesaurus, 2010.

DAHLBERG, Ingetraut. Teoria do conceito. **Ciência da Informação**, Rio de Janeiro, 7(2), p. 101-107, 1978. Disponível em: <http://revista.ibict.br/ciinf/article/view/115/115>. Acesso em: 13 dez. 2020.

FELIPE, Carla Beatriz Marques. **Os aspectos sociocognitivos para a indexação de fotografias**. 2016. Dissertação (Mestrado em Ciência da Informação). Universidade Federal de Pernambuco, Recife, 2016.

FERNEDA, Edberto; DIAS, Guilherme Ataíde. OntoSmart: um modelo de recuperação de informação baseado em ontologia. **Perspectivas em Ciência da Informação**, v.22, n.2, p.170-187, abr./jun. 2017.

FUJITA, Mariângela Spotti Lopes; SANTOS, Luciana Beatriz Piovezan dos; ALVES, Roberta Vesu. Linguagem de indexação e linguagem documentária são sistemas de organização do conhecimento? Uma análise bardiana da variação terminológica. **Scire**, v. 24, n.2, 2018. p. 23-33.

FUJITA, Mariângela Spotti Lopes; RUBI, Milena Polsinelli. Compatibilidade entre linguagens documentais para construção, atualização e adequação de vocabulários de bibliotecas universitárias. *In*: CONGRESO ISKO ESPAÑA Y II CONGRESO ISKO ESPAÑA-PORTUGAL, 12, 2015, Murcia. **Anais [...]**. Murcia: Universidad de Murcia, 2015.

HARPRING, Patricia. **Vocabulários Controlados Terminologia para arte, arquitetura e outras obras culturais**. Tradução de: Christina Maria Müller. São Paulo: Secretaria da Cultura do Estado: Pinacoteca de São Paulo: ACAM Portinari, 2016.

HJØRLAND, Birger. What is knowledge organization (KO)? **Knowledge Organization**, v.35, n.2, 2008. Disponível em: <https://www.researchgate>.

net/publication/277803483\_What\_is\_Knowledge\_Organization\_KO/link/55d8232608aed6a199a6afce/download. Acesso em: 13 dez. 2020.

KOBASHI, Nair Yumiko. **Vocabulário controlado**: estrutura e utilização. Brasília. Escola Nacional de Administração Pública. 2008. Disponível em: <https://repositorio.enap.gov.br/bitstream/1/1289/41/Vocabul%C3%A1rio%20controlado%20-%20estrutura%20e%20utiliza%C3%A7%C3%A3o.pdf> Acesso em: 20 dez. 2020.

MIRANDA, Marcos Luiz Cavalcanti. **Organização e representação do conhecimento: fundamentos teóricos-metodológicos na busca e recuperação da informação em ambientes virtuais**. 2005. 354 f. Tese (Doutorado em Ciência da Informação) - Programa de Pós-Graduação em Ciência da Informação do Convênio CNPq/IBICT – UFRJ/ECO, Rio de Janeiro, 2005. Disponível em: < [tede.dep.ibict.br/handle/tde/54](http://tede.dep.ibict.br/handle/tde/54) >. Acesso em: 02 jan. 2020.

MIRANDA, Marcos Luiz Cavalcanti de; SILVA, Fábio Gomes da. O Islamismo na CDD e CDU: religião e cultura periféricas nos esquemas de classificação bibliográfica. In: BARROS, Thiago Henrique Bragato Barros; TOGNOLI, Natalia Bolfarini (Orgs.). **Organização do Conhecimento Responsável**: Promovendo Sociedades Democráticas e Inclusivas. Belém: Ed. da UFPA, 2019. p. 531-549. (Estudos Avançados em Organização do Conhecimento; 5). Disponível em: <https://isko.org.br/wp-content/uploads/2021/02/LIVRO-ISKO-BRASIL-EDICAO-BELEM.pdf>. Acesso em: 01 set. 2021.

MOREIRO GONZÁLEZ, José Antonio. **Linguagens documentárias e vocabulários semânticos para a web**. Salvador: EDUFBA, 2011.

NATIONAL INFORMATION STANDARDS ORGANIZATION (U.S.). **Guidelines for the construction, format and management of monolingual controlled vocabularies** ANSI/NISO Z39.19 – 2005 (revision of Z39.19-1980). Bethesda (USA): NISO, 2005. 176 p.

NASCIMENTO, Felipe Mozart de Santana; PINHO, Fabio Assis. Sistemas de organização do conhecimento: semelhanças e diferenças. **ConCI**: Convergência em Ciência da Informação, v. 2, n. 3, p. 104-122, set./dez. 2019. Disponível em: <https://brapci.inf.br/index.php/res/v/140503>. Acesso em: 05 dez. 2020.

PINHEIRO, Lena Vania Ribeiro; FERREZ, Helena Dodd. **Tesouro Brasileiro de Ciência da Informação**. Rio de Janeiro; Brasília: Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), 2014. Disponível em: <https://ibict.br/>

images/internas/TESAURO-COMPLETO-FINAL-COM-CAPA-\_24102014.pdf  
Acesso em: 27 ago. 2021

SMIRAGLIA, Richard P. Organización del conocimiento: algunas tendencias em um dominio emergente. **El profesional de la información**, v. 21, n. 3. maio/jun. 2012. Disponível em: <https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/epi.2012.may.01/17913>. Acesso em: 20 jan. 21.

SOUZA, Rosali Fernandez de. Organização do Conhecimento. *In*: TOUTAIN, Lídia Maria Batista Brandão (Org.). **Para entender a ciência da informação**. Salvador : EDUFBA, 2007.

TÁLAMO, Maria de Fátima G. Moreira; LARA, Marilda Lopes, Ginez de; KOBASHI, Nair Yumiko. Contribuição da terminologia para a elaboração de tesouros. **Ciência da Informação**, v.21, n.3, p.197-200, 1992.

TÔRRES, Lecy Maria Caldas. Sistematização da Sintaxe de Cabeçalho de Assunto. c2021. Disponível em: <http://eoci.uff.br/sistematizacao-da-sintaxe-de-cabecalho-de-assunto/>. Acesso em: 01 set. 2021.

ZENG, Marcia Lei. Knowledge Organization Systems (KOS). **Knowledge Organization**, v. 35, n. 2-3, 2008.



# Os Sistemas de Organização do Conhecimento do Repositório de Acesso Livre à Informação Científica (Alice) da Empresa Brasileira de Pesquisa Agropecuária (Embrapa)

Alessa Fabíola dos Santos Ceslinski<sup>1</sup>, Fabiana Magno de Lacerda<sup>2</sup>,  
Lorrane de Souza Saluzi<sup>3</sup>, Maria Cecília Jardim Barros<sup>4</sup>, Skrol Salustiano<sup>5</sup>

---

## 1 Introdução

A EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA (EMBRAPA)<sup>6</sup> É UMA EMPRESA de inovação tecnológica focada na geração de conhecimento e tecnologia para agropecuária brasileira. Vinculada ao Ministério da Agricultura, Pecuária e Abastecimento (Mapa), criada em 26 de abril de 1973.

Em 1974, a empresa deu início às iniciativas relacionadas à organização e disponibilização de informações técnico-científicas por meio de produtos e serviços oferecidos por suas bibliotecas e se dedicou a organizar, disseminar e dar acesso a

---

1 Doutoranda em Ciência da Informação pelo Programa de Pós-Graduação em Ciência da Informação pelo IBICT ECO-UFRJ. Mestre em Gestão de Unidades de Informação pelo Programa de Pós-Graduação em Gestão da Informação (PPGInfo) da Universidade do Estado de Santa Catarina (UDESC). Graduada em Biblioteconomia com Habilitação em Gestão da Informação pela UDESC. E-mail: alessafds@gmail.com

2 Mestranda em Ciência da Informação pelo Programa de Pós-graduação em Ciência da Informação pelo IBICT ECO-UFRJ. Especialista em Gestão Estratégica da Informação pela Escola Politécnica da UFRJ. Graduada em Arquivologia e Biblioteconomia e Documentação pela Universidade Federal Fluminense (UFF). E-mail: fabiana.magnanima@gmail.com

3 Mestre em Ciência da Informação pelo Programa de Pós-graduação em Ciência da Informação pelo IBICT ECO-UFRJ. Graduada em Biblioteconomia e Gestão de Unidade de Informação pela UFRJ. E-mail: lorrannessaluzi@gmail.com

4 Mestre em Ciência da Informação pelo Programa de Pós-graduação em Ciência da Informação IBICT ECO-UFRJ. Graduada em Biblioteconomia e Gestão de Unidade de Informação pela UFRJ. E-mail: mcecilia.jbarros@gmail.com

5 Doutorando e Mestre em Ciência da Informação pelo Programa de Pós-Graduação em Ciência da Informação pelo IBICT ECO-UFRJ. Especialista em Gestão da Comunicação Institucional (UCB). Graduado em Comunicação Social (FAESA)

6 <https://www.embrapa.br/quem-somos>

informações científicas e tecnológicas que permeiam essa estrutura complexa e extremamente produtiva. Nesse contexto, a empresa desenvolveu o Repositório Institucional de Acesso Livre à Informação Científica (Alice), cujo propósito central é dar visibilidade às produções técnico-científicas dos centros de pesquisa da Embrapa.

O repositório da Embrapa Alice é o quarto maior centro de armazenamento institucional de dados do país. Atualmente reúne pesquisas e informações científicas produzidas por pesquisadores da Embrapa, tais como capítulos de livros, artigos em periódicos indexados, artigos em anais de congressos, teses e dissertações, notas técnicas, entre outros tipos de publicações. Apesar de seu acervo ser majoritariamente em língua portuguesa, possui também, documentos em inglês e espanhol. Aliado a uma política de governança de dados, o repositório Alice disponibiliza acesso aberto a todas as suas publicações, baseado em protocolos internacionais, assim como: a Convenção de Santa Fé de 1999, *Open Archives Initiative* (OAI) e e-prints; e a Declaração de Budapeste (*Budapest Open Access Initiative*) de 2002.

Dessa forma, o presente trabalho pretende avaliar os Sistemas de Organização do Conhecimento (soc) do Repositório de Acesso Livre à Informação Científica (Alice) da Empresa Brasileira de Pesquisa Agropecuária (Embrapa), utilizando-se dos métodos de observação direta e análise documentária.

O objeto principal da pesquisa, está na identificação dos soc utilizados pelo Repositório Alice e analisá-los sob a ótica da análise documentária. A avaliação dos soc identificados, atenderá a critérios de indicadores de desempenho do sistema de organização do conhecimento, de modo que possa expressar as expectativas, tanto do classificacionista, quanto do usuário.

### **1.1 Objetivos**

A partir das considerações iniciais, propõe-se como o objetivo geral analisar os SOCs adotados pela Embrapa, com foco no repositório institucional Alice. De forma mais específica, pretende-se identificar quais os soc utilizados pelo Alice e suas tipologias, além de identificar quais suas funções, contexto e componentes.

### **1.2 Métodos**

Com relação aos procedimentos metodológicos adotados, trata-se de uma pesquisa de caráter básico. De acordo com Gerhardt e Silveira (2009, p.34), a pesquisa básica “objetiva gerar conhecimentos novos, úteis para o avanço da Ciência, sem aplicação prática prevista”.

No que se refere à abordagem, a pesquisa se caracteriza como qualitativa. Para Marconi e Lakatos (2010), a pesquisa qualitativa tem por finalidade interpretar e analisar fenômenos, fornecendo uma análise mais detalhada dos mesmos.

Para coleta e análise de dados foi utilizado o método de observação direta do repositório e tesouro, que consiste na coleta direta das informações, sem que haja intervenção dos sujeitos observados. Quanto ao universo de usuários, pode-se afirmar que o mesmo é composto por técnicos, estagiários e pesquisadores da Embrapa, de acordo com Zanelli e Catarino (2014, p.12).

A observação direta tem como suporte um guia constituído a partir de indicadores que designam os comportamentos a se observar. De acordo com Quivy (1992), o processo de observação pode ser segmentado 3 em etapas:

- a) A primeira etapa consiste em conceber um instrumento capaz de produzir todas as informações adequadas e necessárias para testar as hipóteses;
- b) A segunda operação consiste no teste do instrumento de observação;
- c) A terceira é a recolha dos dados.

Quanto ao método utilizado para a coleta de dados, optou-se pela análise documentária. A análise documentária pode ser conceituada como uma atividade metodológica específica no interior da documentação, que trata da análise, síntese e representação da informação” (CINTRA et al., 1994, p.24).

Gardin (1973) apresenta a análise documentária como um conjunto de procedimentos efetuados com a finalidade de expressar o conteúdo dos documentos científicos, sob formas destinadas a facilitar a recuperação da informação. Para García Gutierrez (1990), a análise documentária consiste em extrair as ideias centrais de um documento com o fim de disponibilizá-lo para sua recuperação mediante representações sintéticas.

Cabe ressaltar que uma abordagem qualitativa em pesquisas de âmbito exploratório, tem como principal objetivo, não somente permitir ao pesquisador maior proximidade com o problema de pesquisa, bem como refinar a questão de pesquisa; proporcionar entendimento mais profundo do problema, além de identificar informações que possam ser reunidas para formular as investigativas. Para Marconi e Lakatos (2010, p.141):

O desenvolvimento de pesquisas exploratórias é útil, muitas vezes, para analisar se a estratégia de pesquisa adotada é apropriada e examinar as potenciais abordagens de investigação. A pesquisa exploratória ajuda a verificar a viabilidade do que é proposto em termos de tempo, esforços e recursos, assim como a aplicação adequada e equilibrada desses recursos. Também contribui para gerar ideias iniciais – insights – sobre o fenômeno estudado, assim como para testar a viabilidade de se realizar um estudo mais abrangente.

Para tanto, foi realizado um levantamento bibliográfico dos principais textos que tratam da temática nas bases de dados. Foram utilizados os termos: Alice; Embrapa; Ainfo e repositório, combinados entre si por meio de operadores booleanos. Com relação às bases de dados especializadas em Biblioteconomia e Ciência da Informação nos referidos assuntos, foram elencadas as bases: E-LIS (*E-prints in Library and Information Science*); Brapci - Base de Dados em Ciência da Informação; LISA (*Library and Information Science Abstracts*) e Scielo (*Scientific Electronic Library Online*). Estas bases de dados, embora tenham como principal foco as áreas da Biblioteconomia e Ciência da Informação, abrangem outras áreas do conhecimento uma vez que os periódicos indexados são interdisciplinares. O produto desta pesquisa resultou em: 6 documentos extraídos do próprio repositório Alice; 12 documentos extraídos das referidas bases de dados.

Tal volume documental permitiu o fichamento dos documentos com a finalidade de analisar e sintetizar as informações extraídas, de modo a recolher dados que evidenciem características do sistema de organização do conhecimento utilizado no repositório Alice, são elas:

- a) Características do sistema;
- b) Aplicação do sistema no público alvo;
- c) Proposta inicial - visão atual;
- d) Variáveis, ex: tecnologia;
- e) Forma de classificação;
- f) Qual sistema de classificação;
- g) Critérios avaliativos.

Cabe ressaltar que os métodos aplicados para identificar os devidos critérios para a descrição dos SOCs estão baseados em uma metodologia descritiva com a finalidade de avaliar o sistema por meio de uma observação operacional. Ou seja, a aplicação de técnicas padronizadas, a fim de interpretar a coletar dados por meio da observação objetivando responder questões de comportamento e percepções do usuário. Neste caso as técnicas empregadas referem-se as tarefas do usuário instituída pela IFLA (2016): encontrar; identificar; selecionar; obter, objetivando a recuperação da informação.

## **2 Referencial teórico**

A organização do conhecimento para Dahlberg (2006, p.12), é a Ciência que ordena a estruturação e sistematização dos conceitos (elementos de conhecimento, unidades de conhecimento, etc.), conforme suas características, em classes de

conceitos organizadas em objetos/assuntos, “a fim de permitir uma visão sobre as relações existentes entre eles para que todos possam reconhecê-los e tirar conclusões úteis deles” (DAHLBERG, 2006, p.12) .

A opinião é complementada por Brascher e Café (2008, p. 8) ao avaliar que a Organização do Conhecimento é:

[...] processo de modelagem do conhecimento que visa a construção de representações do conhecimento. Esse processo tem por base a análise do conceito e de suas características para o estabelecimento da posição que cada conceito ocupa num determinado domínio, bem como das suas relações com os demais conceitos que compõem esse sistema nocional.

Embora os critérios para a organização do conhecimento possam variar, a estruturação das relações nela inserida são norteadas conforme determinada definição, implícita ou explícita (MACULAN; DIAS; LARA, 2020, p. 227). Sendo “entidades compostas por unidades de conhecimento organizadas em uma estrutura coesa e adequadamente planejada [...]” (DAHLBERG, 2006, p.12).

A Organização do Conhecimento compreende os Sistemas de Organização do Conhecimento, que de acordo com Carlan e Medeiros (2011, p. 54), são estruturas conceituais “[...] que contemplam termos, definições, relacionamentos e propriedades dos conceitos [...] e cumprem o objetivo de padronização terminológica para facilitar e orientar a indexação e os usuários”. Além disso, afirmam:

No contexto da Ciência da Informação, os SOC ou esquemas de representação do conhecimento são instrumentos que fazem a tradução dos conteúdos dos documentos originais e completos, para um esquema estruturado sistematicamente, que representa esse conteúdo, com a finalidade principal de organizar a informação e o conhecimento e, conseqüentemente, facilitar a recuperação das informações contidas nos documentos (CARLAN; MEDEIROS, 2011, p.55).

Skos (2004, apud Brascher e Carlan, 2010, p. 154) expõe que “Sistema de Organização do Conhecimento é um conjunto de elementos, geralmente estruturado e controlado, que pode ser usado para descrever (indexar) objetos, navegar em coleções, etc”. Além disso, “são instrumentos que dependem de padronização para melhor interagirem, integrarem e interoperarem entre eles e entre humanos, sejam usuários, provedores de informações ou máquinas” (CARLAN; MEDEIROS, 2011, p. 72)

Com o desenvolvimento da internet, os soc têm sido utilizados de forma automatizada, “[...] para o tratamento e a recuperação da informação, tanto no ambiente web como no tradicional, inclusive no desenvolvimento da web semântica” (CARLAN; MEDEIROS, 2011, p.56)

Independente do ambiente, os soc cumprem duas funções principais: na organização da informação e do conhecimento, tratando da padronização da representação dos conteúdos dos documentos e, na recuperação da informação, orientando e auxiliando o usuário quanto à localização de conteúdo. E que incluem uma variedade de esquemas como os tesouros, classificações, listas de cabeçalhos de assunto, terminologias, taxonomias e ontologias. Assim, Brascher e Carlan (2010, p. 173) apontam:

Os SOC's apresentam pontos comuns: compartilham princípios de classificação; utilizam algum tipo de signo para representar os conceitos; cumprem a função de padronizar os termos empregados para representar os conceitos; apresentam recursos para delimitar o significado e estabelecem relações entre conceitos. Observamos, no entanto, que cada sistema possui peculiaridades quanto à estrutura e funções.

Porém, não existe modelo padrão para descrever os soc, dessa forma vamos utilizar a proposta de Zeng (2009), por estar alinhada aos objetivos da pesquisa e apresentar uma concepção baseada na complexidade de suas estruturas e principais funções.

Os SOC's são modelos de representação que auxiliam o profissional da informação a estabelecer mecanismos de indexação, contribuindo para o controle terminológico do assunto. Os soc fornecem a especificidade da terminologia e delimitam junto à comunidade inserida padrões de identificação entre as informações. Além de contribuir para estruturação, a classificação, delinear relações e conceitos de uma área de interesse. Estes fatores ajudam a controlar sinônimos, diminuir o índice de ambiguidade, fazer a relação dos termos de forma hierárquica ou associativa e definir conceitos.

Embora os SOC's, como Tesouro, Classificação, Taxonomia e Ontologia, busquem a padronização de uma estrutura sistemática para a definição da terminologia, seus níveis de representação variam conforme a especificidade do tipo de Sistema de Organização do Conhecimento. A classificação apresenta o código numérico e o termo específico, o Tesouro e a taxonomia termos puros consolidados na área, enquanto a ontologia utiliza os termos de buscando a compatibilidade com outros padrões como o XML (eXtensible Markup Language), RDF (Resource Description Framework) e OWL (Web Ontology Language).

Seguindo a trajetória de apresentação dos SOCs realizada por Carlan e Medeiros

[...] deu-se destaque aos tesouros, pelo uso ainda frequente na área de organização de informação e por sua existência consolidada e padronizada por normas internacionais (ISO, ANSI/NISO); as taxonomias, por sua importância na organização de informações em empresas e instituições, principalmente, no desenvolvimento de portais no ambiente web; as ontologias, pelo interesse da comunidade de pesquisa na área, em função das promessas da web semântica e, ainda pelo potencial que oferecem em relação à capacidade de representação do conhecimento de forma complexa e completa e os sistemas de classificação, pela ampla utilização na organização da informação em bibliotecas, onde são empregadas até hoje. (CARLAN; MEDEIROS, 2011, p. 55)

O termo tesouro é originário do latim – *thesaurus* e do grego – *thesaurós*, que quer dizer tesouro, armazenamento, repositório (BRASCHER E CARLAN, 2010, p. 158). De acordo com Gomes (1996), consistem em “um conjunto de termos semântica e genericamente relacionados, cobrindo uma área específica do conhecimento. É um instrumento da indexação/recuperação de informação” e se apresenta como um instrumento alfabético de controle, onde cada termo possui relação, seja lógica ou semântica com outros termos.

Para Pinheiro e Ferrez (2014) o tesouro, instrumento essencial para consistência de terminologia e de vocabulário de dado campo do conhecimento, “tem ampla aplicação não somente para indexadores, como também pesquisadores, professores e profissionais de informação em geral”.

Sales e Café (2011, p. 102) expõem:

Tesouros são vocabulários controlados formados por termos-descritores semanticamente relacionados, e atuam como instrumentos de controle terminológico. Os tesouros podem estar estruturados hierarquicamente (gênero-espécie e todo-parte) e associativamente (aproximação semântica), e são utilizados principalmente para indexar e recuperar informações por meio de seu conteúdo.

A organização dos tesouros se estrutura em duas bases: teórica, que é a essência e a fundamentação para a construção de tesouros: conceito, termo, categorias e faceta; e técnico-operacional, abordagem mais prática no desenvolvimento de tesouros: planejamento, coleta de termo, controle terminológico, estabeleci-

mento de relações entre conceitos e formas de divulgação e publicação (CARLAN; MEDEIROS, 2011, p.56).

Portanto, tesouros são um conjunto de vocábulos cujos descritores possuem o objetivo de representar os termos, para expressar os conceitos de relevância em um domínio, podendo ser elaborado tanto para áreas mais gerais ou para áreas mais específicas, uma vez que são alicerçados através uma rede lógico-semântica (MACULAN; DIAS; LARA, 2020, p. 198).

Taxonomias são instrumentos que organizam logicamente os conteúdos informacionais. O conceito de organização remete para um procedimento classificatório e permite agrupamento categorizado, isto é, a partir de um assunto formam-se categorias que se dividem em classes e subclasses hierarquicamente, formando uma lista de categorias de assunto estruturada. As taxonomias permitem que se estabeleçam padrões de classificação e ordenação de informações por meio de herança, ou seja, pelo relacionamento hierárquico (gênero/espécie) entre os objetos, em que as características das classes gerais são repassadas por herança às subclasses. Na organização e recuperação de informação, as taxonomias são usadas para a criação de metadados ou termos para descrever um objeto e na categorização para definir classes e subclasses, como suporte à navegação no ambiente web (BRASCHER E CARLAN, 2010).

Observa-se que as taxonomias adquirem importância no contexto das organizações, como instrumentos auxiliares à gestão do conhecimento e a organização e recuperação da informação. As taxonomias asseguram que todas as aplicações da instituição utilizem a mesma linguagem para organizar, armazenar e apresentar a informação (BRASCHER e CARLAN, 2010, p. 160). Além disso, criam uma rede semântica comum baseada na necessidade dos negócios e consideram os bens intelectuais e a maneira pela qual os funcionários procuram pela informação” (CARLAN; MEDEIROS, 2011, p.57).

As ontologias, assim como os tesouro e taxonomia, derivam do grego - *ontos* e *logoi*, que significa conhecimento do ser. A área de informação toma o termo ‘emprestado’ da Filosofia, em cujo âmbito refere-se à teoria sobre a natureza da existência (BRASCHER e CARLAN, 2010, p. 160).

A ontologia é um modelo de representação do conhecimento que, a exemplo do tesouro, é utilizada para representar e recuperar informação por meio de estruturas conceituais. As ontologias possibilitam compartilhar uma forma de pensar de determinado assunto, proporcionando um mapa semântico e uma estrutura conceitual de um domínio específico por meio de um vocabulário comum, pois:



[...] transcendem esta meta de padronizar a linguagem utilizada na indexação e na recuperação da informação, propondo ser um mapa semântico, uma estrutura formal para um dado domínio ou até mesmo servir como ferramenta capital para a elaboração de bases de conhecimento (SALES E CAFÉ, 2009, p. 111).

O diferencial desta linguagem de indexação é que não visam à ‘tradução’ de linguagens naturais para linguagens especializadas e vice-versa, mas, sim, atuam no próprio processamento dessas linguagens (SALES E CAFÉ, 2009, p. 112). Por este motivo, o uso do termo ontologia tornou-se muito frequente na ciência da computação, principalmente na subárea da representação do conhecimento, sendo um dos principais objetivos desse uso ser a construção de bases de conhecimento interoperáveis e melhor estruturadas (MOREIRA; ALVARENGA; OLIVEIRA, 2020, p. 2).

As classificações bibliográficas tiveram grande parte do seu desenvolvimento no final do século XIX e início do século XX e se originaram dos modelos de classificação filosófica, porém, com o objetivo de organizar e localizar os documentos. São sistemas predeterminados de conceitos logicamente estruturados e acompanhados de um código identificador. Este código é atribuído a conceitos ou a documentos em função da correspondência de assuntos (BRASCHER E CARLAN, 2010, p. 157).

Classificar é ordenar, organizar, reunir segundo características comuns. Na Biblioteconomia, esta organização é um imperativo e ela se dá através de instrumentos como tabelas de classificação – criadas com o fim específico de organizar fisicamente as coleções de documentos. A literatura tem mostrado que a classificação está presente não apenas nos sistemas de recuperação de informação, mas na base de sistemas e atividades que se ocupam da organização do conhecimento em suas diferentes manifestações como, por exemplo, os sistemas de inteligência artificial e hipertextos. A unidade a ser manipulada nestes sistemas e atividades é o conceito, que é uma unidade de conhecimento (GOMES, 1996).

A classificação é, provavelmente, o método mais simples de ordenar a complexa multiplicidade da natureza. É um processo de seleção de ideias ou objetos em grupos, conforme suas qualidades semelhantes e diferenças específicas. Esse processo, chamado de abstração, é essencialmente mental, nós agrupamos ou separamos coisas de acordo com o conceito ou ideia que temos dos objetos no mundo, fortalecendo a memória e o poder de raciocínio. Sem isso, nada pode ser identificado, portanto, na prática, pode-se dizer que os pensamentos e raciocínios consistem de classificação. O processo de classificação é uma formação metodológica e sistemática onde se estabelecem critérios para a divisão, isto é, a formulação de um esquema de categorias, classes e subclasses, baseado nas características e relações

dos objetos considerados. É, também, um sistema logicamente estruturado onde os conceitos pré-determinados correspondem a um código identificador (BRASCHER E CARLAN, 2010, p. 157).

Conforme Cintra et al. (2001, p.21) elucida, as Terminologias fazem parte do escopo das Linguagens Documentárias e, no processo de elaboração das linguagens este escopo é dilatado, abarcando as linguagens de especialidade e linguagem natural. Além disso, a Terminologia como campo inter e transdisciplinar trabalha com a descrição e ordenamento do conhecimento no nível cognitivo, com a transferência do conhecimento em nível comunicacional e com os elementos centrais, conceitos, termos e áreas de experiências no nível de normas internacionais (CARLAN; MEDEIROS, 2011, p.65).

A Terminologia Normativa que mediante a normatização dos termos auxilia no processamento técnico da informação e da recuperação da informação, ao utilizar parâmetros gramaticais e semânticos normativos (CINTRA et al., 2001, p. 21). As terminologias tratam de uma área do conhecimento, que denomina conceitos, ou seja, conjunto de metodologias e diretrizes que regem a descrição, formação dos termos e estruturação de campos conceituais. Refere-se a um conjunto de termos de um campo específico do conhecimento relacionados com uma língua de especialidade, sendo produtos gerados pela prática. Para Moreira (2010, p. 32-33), a terminologia, assim como as ontologias, que são “conjuntos estruturados de conceitos especializados representados por denominações”, pode auxiliar de modo importante a compreensão da descrição.

De acordo com Cintra et al. (2001, p. 21), a Terminologia Descritiva tem por exclusividade a produção de classificações conceituais, onde o processo de construção de linguagens documentárias necessita dos recursos desses instrumentos terminológicos.

Sobre a Teoria Geral de Terminologia (TGT), Carlan e Medeiros apontam:

A TGT, pioneira e responsável pela base do trabalho terminológico, possui uma perspectiva prescritiva e normativa que visa rotular e padronizar definitivamente o uso de termos e conceitos, sem considerar as inúmeras variações possíveis em um contexto comunicacional. A Teoria da Socioterminologia de François Gaudin, sob uma perspectiva descritiva com ênfase ao uso social da língua, aceita a variação e a flexibilidade lexical e conceitual, proporcionando uma aproximação entre mecanismos de informação e usuários. E a Teoria Comunicativa da Terminologia de Maria Teresa Cabré fundamenta aspectos comunicativos das línguas naturais para melhor conduzir a comunicação entre especialistas”.(CARLAN ; MEDEIROS, 2011, p.65)

De acordo com Cabré (1993, p. 169), sobre uma área específica do conhecimento, a terminologia é constituída pelo conjunto de palavras especializadas de uma determinada disciplina. Os termos, que são as unidades da base da terminologia, designam os conceitos próprios de cada disciplina especializada.

Cabeçalhos de assuntos são linguagens documentárias estruturadas e pré-coordenadas, oferecendo ao indexador possibilidades limitadas de representar os documentos. Oferecem organização enumerativa, não hierárquica e, via de regra, o arranjo é ordenado de forma alfabética. Como apresentam o princípio de pré-coordenação, representam um vocabulário controlado onde os termos devem ser escolhidos em uma lista pré-existente, sendo, portanto, sistemas fechados de recuperação da informação, exercendo uma função prescritiva (CESARINO E PINTO, 1978).

Por fim, compreende-se que alguns dos SOC apresentados podem ser caracterizados como linguagens documentárias e vocabulários controlados, possuindo especificidades e diferenças entre si. O processo de escolha do SOC adequado está diretamente ligado ao objetivo dos processos de indexação envolvidos.

A partir do referencial teórico foi possível o entendimento e compreensão dos conceitos, possibilitando identificar, dentre as questões propostas nesta pesquisa, se o SOC a ser pesquisado se tratava especificamente de um tesouro.

### **3 Sistemas de Organização do Conhecimento (SOC), subsistemas e avaliação**

O Sistema Embrapa de Bibliotecas (SEB)<sup>7</sup> possui uma coordenação e 43 bibliotecas distribuídas nas Unidades da Embrapa, em todo o Brasil, cujo acervo possui, aproximadamente, 900 mil itens (diferentes tipologias, mídias e formatos). Atualmente, está vinculado à Gerência de Comunicação e Informação da Secretaria-Geral da Embrapa, que agrega as bibliotecas em nível técnico e procura contribuir para o trabalho colaborativo e padronizado dos profissionais que nela atuam (BELTRÃO et al., 2017, p.123).

O projeto começou no ano de 1974, com as primeiras iniciativas em relação à organização e disponibilização de informações técnico-científicas quando foi estruturado o Sistema de Informação Técnico-Científica da Embrapa (SITCE), criado o Departamento de Informação e Documentação (DID) e bibliotecas em diversas Unidades Descentralizadas.

Em 1975, teve início o processo de centralização das atividades de aquisição e catalogação para o DID. Em 1976, foram criados novos serviços relacionados à informação técnico-científica, como: comutação bibliográfica, disseminação seletiva

7 <https://www.embrapa.br/seb/o-sistema>

da informação, elaboração de resumos informativos, elaboração de bibliografias especializadas e apoio na editoração. Neste ano, também aconteceu o I Encontro de Bibliotecários da Embrapa, o DID firmou convênio com a *National Agricultural Library* dos Estados Unidos e com a *British Library*, o objetivo era automatizar a base de dados das bibliotecas a empresa. Em 1977, o DID foi transferido para o Edifício Venâncio 2000 e teve início a elaboração de bibliografias e catálogos em conjunto com outros países como Chile, Bolívia, Argentina e França.

Em 1982, aconteceu o lançamento do primeiro Catálogo Coletivo de Periódicos da Embrapa (CCPE) e o primeiro Curso de Especialização em Informação Agrícola (IBICT/BINAGRI/Embrapa e CBDA). O DID foi extinto em 1983 e suas atribuições passaram a ser de competência do Departamento de Difusão de Tecnologia (DDT). O DID voltou a ser criado em 1986. No ano de 1988, houve a operacionalização do Sistema Brasileiro de Informação sobre Pesquisas Agrícolas em Andamento (BRACARIS) e a Inauguração do edifício da Embrapa Sede. Em 1989, há o investimento e a aquisição de potentes computadores para a informatização dos acervos das bibliotecas.

O DID foi novamente extinto no ano de 1990 e foi criado o Departamento de Informação e Editoração (DIE). Porém, o DIE foi extinto em 1991 e substituído pelo Serviço de Produção de Informação (SPI), estrutura descentralizada da Embrapa.

Em 1991, ocorreu a descentralização dos serviços de comutação bibliográfica nacional e internacional, a desativação dos serviços de intercâmbio central, banco de bibliografias, banco de teses e banco de tradução. Para atender as novas demandas foi criado o Conselho Técnico Consultivo para as Áreas de Informação, implantado o Sistema Embrapa de Informação (SEI), com enfoque sistêmico da informação e lançada a primeira versão do software Ainfo – sistema informatizado para gestão do acervo das bibliotecas, pela Embrapa Informática Agropecuária.

Com a tecnologia se tornando parte fundamental do processo, em 1993, as atividades de informação e documentação passaram a ser de competência da Coordenadoria de Informação e Documentação (CID), vinculada ao Departamento de Informática (DIN), que, no mesmo ano, passou a ser denominado de Departamento de Informação e Informática (DIN).

No ano de 1998, foi lançada a Base de Dados da Pesquisa Agropecuária, em CD-ROM, com todo o catálogo das bibliotecas da Embrapa, ocorreu a aprovação do Manual de Informação e Documentação pela Diretoria Executiva. Em 1999, foi implementado o serviço de alerta eletrônico, aplicativo para divulgar via intranet, os sumários de periódicos assinados pela Embrapa.

Em 2003, a Embrapa firmou convênio com a Capes e passou a ter acesso ao Portal de Periódicos da Capes, via usuário colaborador, a CID tem suas atribuições transferidas para a Embrapa Informação Tecnológica, unidade que passa a geren-

ciar o SEB. Em 2005, houve participação do SEB no Conselho Técnico do Programa COMUT. Enquanto a aprovação pelo Sistema Embrapa de Gestão do Projeto Memória Embrapa e do Projeto Desenvolvimento e Gestão do SEB ocorreram em 2006.

No ano de 2007, houve a aquisição de e-books de importantes editores internacionais e a integração do Sistema de Informação de Apoio à Decisão Estratégica (SIDE). Em 2008, foi lançada a Biblioteca Digital e aprovado o Projeto Acesso Aberto na Embrapa. Em 2009, aconteceu o I Workshop de Gestão da Informação da Embrapa.

Em 2011, a EMBRAPA lança os repositórios institucionais Alice e Infoteca-e, além do metabuscador Sabiia. Em 2013, foi realizado o Seminário Tendências da Gestão da Informação em Instituições de Ciência e Tecnologia, pela Embrapa Informação Tecnológica, em parceria com a Embrapa Estudos e Capacitação. No ano seguinte, 2014, ocorreu a aprovação do Manual de Procedimentos para Sistematização, Desenvolvimento e o Registro de Memória Oral da Embrapa.

Em 2014, a equipe do Ainfo capacitou a equipe do Instituto Nacional de Investigação Agropecuária para desenvolvimento e implantação do Sistema Ainfo 6. Em 2017, foi lançada a segunda versão dos repositórios digitais de acesso aberto à informação científica e tecnológica: Alice e Infoteca-e, com a adoção da licença pública de uso *Creative Commons*, é aprovado pelo Sistema Embrapa de Gestão o arranjo Gestão da Informação Científica e Tecnológica da Embrapa: estratégias para o fortalecimento, consolidação e inovação (GESTINF). Em 2018, o SEB se tornou vinculado ao setor de Gestão da Informação, da Gerência de Comunicação e Informação, da Secretaria Geral da Embrapa.

### 3.1 Alice

Lançado em Abril de 2011, o Repositório Institucional de Acesso Livre à Informação Científica da Embrapa, conhecido como Alice, tem como objetivo a expansão da visibilidade da produção técnico-científica dos centros de pesquisa da Embrapa, bem como servir de indicador tangível da sua qualidade com intuito de reunir, armazenar, organizar, recuperar, preservar e permitir o acesso ao texto completo, assim como capítulos de livros, artigos em periódicos indexados, artigos em anais de congressos, teses e dissertações, notas técnicas, entre outros.

O Alice foi desenvolvido pelo SEB, sob a responsabilidade da Embrapa Informação Tecnológica, e construído utilizando o software DSpace (figura 1).

Figura 1 – Homepage do Repositório institucional Alice



Fonte: <https://www.alice.cnptia.embrapa.br/>. Acesso em 29 out. 2020.

Atualmente, o Alice conta com 101.698 itens depositados e apesar da visibilidade em mais de 244 países, o volume de downloads e consultas estão concentrados em apenas quatro países: Estados Unidos, Brasil, Alemanha e China. Entre as funcionalidades do repositório é possível visualizar as estatísticas geradas pela página do próprio repositório. Dentre elas pode-se citar: volume documental de downloads e consultas por período e país, itens e coleções que foram mais consultados ou que tiveram o maior número de download e algumas estatísticas administrativas. Durante o processo de representação documental, o bibliotecário faz o *upload* do arquivo em formato PDF e inclui na coleção digital a que este documento se enquadra conforme sua natureza científica ou tecnológica.

O depósito dos itens que constituem o Alice é feito pelos bibliotecários da Embrapa, encarregados da catalogação, indexação e carregamento do arquivo em formato digital no sistema informatizado para gestão de acervos impressos e digitais de bibliotecas (Ainfo). Entretanto, o repositório também conta com uma política de auto arquivamento, permitindo aos próprios usuários, enquanto autores dos trabalhos, o depósito de sua produção científica. Tal fato acarreta, infelizmente, numa certa falta de padronização na descrição dos metadados, o que afeta diretamente as pesquisas no repositório, podendo gerar resultados inconsistentes.

Quanto a disponibilização dos documentos no repositório, Castro afirma que:

[...] é importante ressaltar que toda a produção científica dos pesquisadores deve ser registrada, mas nem todos os documentos podem ser disponibilizados na íntegra nos repositórios. Os documentos editados

pela Embrapa que sejam para publicação em acesso aberto são disponibilizados na íntegra nos repositórios. Já os artigos publicados em periódicos que não são de acesso aberto ou em livros e capítulos de livros dos quais a Empresa não disponha de termo de cessão de direitos patrimoniais assinado pelos autores, são depositados nos repositórios com restrição para disponibilização na íntegra (CASTRO, 2013, p.13).

Essa ressalva apontada por Castro (2013) acontece devido o sistema de bibliotecas ser regido por Leis, Decretos, Portarias, Instruções Normativas, Normas Complementares, Resoluções e Atos Administrativos, que atualmente totalizam 52 especificações de procedimentos, para regular como os documentos produzidos e sobre a guarda da Embrapa poderão ser acessados pelo público. Contudo, cabe ressaltar que, apesar de nem todos os documentos possam ser acessados livremente, os metadados que o descrevem se encontram em acesso aberto, permitindo assim a conexão com outros documentos.

Dentro deste conjunto de documentos que norteiam a implantação da política de dados, como serão armazenados, catalogados e disponibilizados, é possível observar a forte influência das Políticas de Governo. Com legislações e atos gerando um emaranhado sistema para a divulgação científica da Instituição.

### 3.2 Thesagro

O Thesaurus Agrícola Nacional (THESAGRO) é o único tesouro em português brasileiro especializado em literatura agrícola. Foi criado e desenvolvido segundo diretrizes da UNESCO, estabelecidas pela *United Nations Information System (UNISIST)*, por meio de documento *Principes directeurs pour l'établissement et le développement de thesaurus monolingues* (SC/WS/555, Paris, 1973).

O Thesagro foi desenvolvido pela Coordenação Geral de Informação Documental (Cenagri). Posteriormente, com as alterações do organograma do Ministério da Agricultura a responsabilidade de guarda e manutenção foi transferida para a Biblioteca Nacional de Agricultura (BINAGRI), que oferece suporte informacional a todos os órgãos do Ministério da Agricultura, Pecuária e Abastecimento (MAPA). Entre as responsabilidades da Biblioteca Nacional de Agricultura (BINAGRI) está a responsabilidade pela coleta, padronização e tratamento de informações disponíveis em bases de dados e outras fontes, atividades que facilitam a recuperação e disseminação de dados do domínio da agricultura e de áreas correlatas.

A primeira edição do THESAGRO foi publicada em junho de 1979, com o nome “Thesaurus para Indexação/Recuperação da Literatura Agrícola Brasileira”, desenvolvido pela antiga Coordenação de Informação Documental Agrícola - CENAGRI.

Em 1989, houve a publicação de uma nova versão revisada e, depois, somente foi atualizado em 2006. Atualmente, é disponibilizado no formato *Extensible Markup Language* (XML), para possibilitar a utilização gratuita.

O objetivo do THESAGRO é, principalmente, a recuperação de documentos e ele se caracteriza pela sua especificidade, possuindo um escopo amplo, cobrindo assuntos desde energia até educação. No final do ano de 2020, o Thesagro possuía mais de nove mil termos. Entretanto não incluindo categorias de assunto, e específica as relações entre termos e conceitos por meio de remissivas: (a) de equivalência USE (“use”; utilize) e UF (“used for”; utilizado para), para indicar sinônimos; (b) hierárquicas BT (“broader term”; termo genérico) e NT (“narrower term”; termo específico); (c) associativas RT (“related term”; termo relacionado).

Uma das características do Thesagro é a utilização de toda a terminologia em português, porém os símbolos aplicados para representar os relacionamentos estão apresentados no idioma inglês.

### 3.3 A expansão das informações agrícolas

Durante a década de 1970 houve uma rápida expansão dos usuários dos serviços de Bibliografias Personalizadas em Agricultura (BIP, AGRI), um serviço de disseminação seletiva da informação desenvolvido pelo Sistema Nacional de Informação e Documentação Agrícola (SNIDA), a necessidade de aumentar o nível de especificidade na recuperação e de trabalhar com arquivos contendo quantidades crescentes de referências, levaram a uma ampliação do sistema inicial de processamento e a incorporação de novos programas especialmente desenvolvido para este fim.

Neste período, a circulação das informações era baseada em um programa de processamento de recuperação que utilizava, como critérios de busca, os códigos do esquema de categorização do Sistema AGRIS ou suas combinações. A descrição geral deste programa, denominado “MEDUSA” (CHASTINET. et al. 1978). Este programa que antecedeu o Thesagro recuperava a informação de acordo com as estratégias de busca representativa do perfil de interesse do usuário. As estratégias de busca eram baseadas em expressões booleanas do tipo and (+) e or (i) de dois ou mais códigos AGRIS.

No entanto, o número de usuários registrava aumento constante, o sistema AGRIS demonstrava a necessidade de um mecanismo de informação mais ágil e abrangente. Além de demonstrar muitas limitações na recuperação das informações. Para sanar esse gargalo foi adaptado ao sistema AGRIS o TEXTEPAC e ampliado o número de base de dados.

Dentro desse cenário, que demandava uma forma de cadastrar registros nacionais de documentos existentes em diversas formas de suportes físicos e eletrônicos e recu-



perar as informações contidas nas várias bases de dados, foi implementado o Thesagro. Que teve a normatização do Ministério da Agricultura somente no ano de 1990, com a Portaria nº 291/1990, seção VI, artigo 120, que preconizou ser uma das funções da Coordenação da Informação e Documentação Agrícola: “elaborar versão, em língua portuguesa, do “Thesaurus/Agrovoc”, desenvolvido pela FAO/Aggris.” (BRASIL, 1990).

### 3.4 Avaliação do Thesagro

Ao realizar a observação direta do Alice, percebeu-se que o mesmo conta com três SOCs: o Thesagro - Thesaurus Agrícola Nacional; o *National Agricultural Library's Thesaurus and Glossary (NAL Thesaurus)* da Biblioteca Nacional de Agricultura dos Estados Unidos; e as palavras-chave do autor.

Para o melhor desenvolvimento deste trabalho, optou-se pela análise do Thesagro, por tratar-se de um tesouro brasileiro da área de agricultura, a fim de que o mesmo possa compor um repositório nacional de SOCs.

O Thesagro foi desenvolvido segundo as normas estabelecidas pela *United Nations Information System (UNISIST)* do documento *Principles directeurs pour L'établissement et le développement the thesaurus monolingues (sc/ws/555, Paris, 1973)*. O vocabulário especializado em agricultura teve sua primeira edição em 1979, Thesaurus para Indexação/Recuperação da Literatura Agrícola Brasileira, e a edição seguinte em 1989. Em 2006, foi lançada a versão virtual do Thesagro - Thesaurus Agrícola Nacional, com o equivalente a 9.763 termos, interligados por três relações básicas: hierárquicas (TG, TE), associativas (TR) e equivalência (USE, UP).

O aperfeiçoamento e o enriquecimento do Thesagro se devem à colaboração de instituições agrícolas brasileiras com a sua desenvolvedora, a Biblioteca Nacional de Agricultura (BINAGRI).

O Domínio do Conhecimento do Sistema de Organização do Conhecimento Thesagro é a Agricultura e é utilizado para o controle terminológico nos processos de indexação e recuperação dos documentos dos produtos da Embrapa, entre eles, o Alice, que faz parte do sistema Ainfo. A interface não apresenta introdução sobre o SOC, mas contém somente termos relacionados à Agricultura. O uso do tesouro é intuitivo, porém para fazer as correlações é preciso ter um conhecimento prévio do assunto. A versão digital não apresenta maiores informações acerca do tesouro, conforme figura 2.

**Figura 2 – Homepage do Thesaurus Agrícola Nacional (THESAGRO)**

Thesaurus Agrícola Nacional

Início Minha conta  Buscar Pesquisa avançada Sobre

Thesaurus Agrícola Nacional / Biblioteca Nacional de Agricultura - BINAGRI

URI <http://sistemas.agricultura.gov.br/tematres/vocab/>  
 Idioma pt  
 Contact mail  
 Data de criação 30/09/2016  
 Data da última alteração 05/06/2020  
 Palavras chave  
 Tipo de linguagem Controlled vocabulary  
 Cobertura  
 Termos 3763 [Ver estatísticas](#)  
 Relações entre termos 12021  
 Termos não preferidos 1154  
 Nota de escopo 91  
 API <http://sistemas.agricultura.gov.br/tematres/vocab/services.php>  
 Criado por TemaTres 2.2

ES 4.0 CC 0 BY NC ND SA TemaTres [português](#)

Fonte: <http://sistemas.agricultura.gov.br/tematres/vocab/sobre.php>. Acesso em 23 out. 2020.

Trata-se de um tesouro monolíngue em português e sua principal função está pautada em auxiliar na classificação e indexação dos materiais bibliográficos na área da Agricultura, sobretudo no que tange o contexto dos repositórios Alice e Infoteca-e da Embrapa. Apresenta componentes da parte sistêmica: relações de hierarquia entre os termos. E componentes da parte alfabética: lista alfabética dos descritores, com relações de ordem lógica, ontológica e de equivalência.

As relações semânticas usadas são: termo específico (TE), termo geral (TG) e termo relacionado (TR). As relações semânticas são cruciais pois são estabelecidas analisando-se as propriedades dos conceitos, as quais nos permitem identificar as diferenças e similaridades que mostram certos tipos de relacionamentos. Além dos elementos USE e UP.

Os termos são apresentados em ordem alfabética e estruturados hierarquicamente. Não apresenta outro tipo de apresentação e faz o uso de termos como unidade linguística. Têm consistência das relações entre os termos e consistência no nível de especificidade, porém não apresenta consistência no uso do plural e do singular, de acordo com a figura 3.

Não foram encontradas notas de aplicação, escopo e nem a política de indexação. Apresenta linguagem lógica e está disponível no endereço eletrônico <http://sistemas.agricultura.gov.br/tematres/vocab/index.php>. O Thesagro na Embrapa é utilizado pelo sistema Ainfo, que alimenta automaticamente os repositórios Alice e Infoteca-e. A versão virtual foi desenvolvida com base no TemaTres, que é um servidor de vocabulário de código aberto, uma aplicação web para gerenciar e explorar vocabulários, tesouros, taxonomias e representações formais do conhecimento.

**Figura 3 – Relações semânticas do Thesagro**

The image displays two side-by-side screenshots of the Thesagro interface. The left screenshot shows the term 'Anu Branco' with its preferred terms (Anum Branco, Anum Branco) and generic terms (Ave Ornamental). The right screenshot shows the term 'Anum Branco' with its preferred terms (Anum Branco, Anum Branco) and generic terms (Ave Ornamental). Both screenshots show a list of terms on the left and a detailed view of the selected term on the right. The detailed view includes metadata such as creation date, modification date, and various statistics.

Fonte: Elaborado pelos autores.

## 5 Considerações finais

O repositório institucional Alice cumpre sua missão ao ampliar a visibilidade da produção técnica e científica da Embrapa, sobretudo no que tange servir como indicador de qualidade das pesquisas em Agricultura, área central da Embrapa, que pode ser observado pelas estatísticas de acesso e downloads.

Formado por uma equipe especializada de profissionais em gestão da informação, os documentos são catalogados e indexados antes de serem incluídos em formato digital no repositório. Contudo, a política de auto arquivamento institucional, onde os próprios autores carregam o arquivo digital, provoca conflitos quanto à constituição de seus metadados e conceitos utilizados para representação documental. A pesquisa não identificou ações, por parte dos profissionais da informação da Embrapa, quanto a curadoria desses documentos de auto arquivamento. Logo, pode-se concluir que apesar da aplicação dos tesouros da área, a política de auto arquivamento não é clara quanto a classificação desses documentos.

A falta de padronização terminológica acarreta em resultados discrepantes para um mesmo conceito, o que se agrava com a ausência das relações de equivalência entre os termos pesquisados. Ademais, apesar de fazer uso de lógica booleana, o usuário do repositório Alice, em sua maioria estagiários de pesquisadores, não o utilizam.

Com relação à análise e avaliação do Thesagro, o tesouro brasileiro de agricultura, concluiu-se que o mesmo atinge seu objetivo de descrever e representar as informações científicas e tecnológicas da Embrapa, além de possuir consistência

de relações e especificidade, já sendo um tesouro bastante consolidado. Entretanto, apresenta algumas falhas como a falta de uma introdução, notas de escopo e aplicação, consistência no uso do plural e singular e uma política de indexação.

Dessa forma, entende-se que uma política de indexação da Embrapa poderia abarcar não somente a diversidade de itens que compõem o acervo do repositório institucional Alice, bem como ao universo de usuários que ela atende. Certos aspectos precisam estar claramente discriminados numa política de indexação para que não haja retrabalho e que se possa fazer uma análise, de modo a utilizar o sistema como um dado de pesquisa no aprimoramento da linguagem. Identificar os pontos fortes e fracos de um sistema de organização do conhecimento, torna-se crucial para a tomada de decisão no desenvolvimento de uma política de indexação. Haja vista que os sistemas de organização do conhecimento são sempre de natureza de pesquisa, nenhum sistema pode ser visto como um instrumento fechado, pois deve acompanhar o desenvolvimento da área de conhecimento.

Cabe ressaltar que o papel do sistema de organização do conhecimento na Embrapa é extremamente difícil, amplo e complexo, tanto do ponto de vista da abrangência quanto da especificidade. A importância da Embrapa, enquanto instituição de pesquisa, se justifica uma vez que o país passa da figura de importador, para exportador. Evidenciando, assim, uma configuração de extrema mudança em relação a todos os aspectos informacionais e tudo que deriva desta área de conhecimento. Pois, lida com questões de ensino, em suas teses e dissertações; a questão da pesquisa num amplo espectro que possa vir a abranger o conceito interdisciplinar da Agricultura; e por fim, a questão da extensão no sentido de aplicar aquilo que está sendo desenvolvido nacional e internacionalmente.

Observou-se que as falhas decorrentes da política de auto arquivamento poderiam ser sanadas por ações de competência informacional, à medida que oriente os depositantes quanto ao termo correto a ser utilizado. Além disso, percebeu-se que o universo de usuários do repositório, apesar de ter certo grau elevado de instrução, faz pouco uso das ferramentas de busca. O uso dos recursos como ferramentas de busca, são instrumentos indispensáveis na recuperação da informação, tanto na pesquisa livre, como na estruturada. O repositório Alice, permite o uso de ferramenta de buscas, assim como: operadores booleanos, truncagem, mascaramento, busca por proximidade e frase, dentre outras possibilidades na formulação de estratégias de busca. Resultando, assim, não somente numa busca exaustiva e alto índice de revocação, bem como em resultados precisos, úteis e relevantes.

Doravante, podemos concluir que o uso dos vocábulos controlados, NAL (*National Agricultural Library*) e Thesagro, especificamente, abrange de forma satisfatória os processos de indexação e recuperação dos documentos inseridos no repo-

sitório. Uma vez que é constituído por diversas instituições agrícolas brasileiras, de modo a viabilizar o desenvolvimento constante de sua terminologia.

### Referências

BELTRÃO, S. L. L.; SILVA, A. R.; DANTAS, J. O.; ARRUDA, R. G. Atuação do Sistema Embrapa de Bibliotecas e a gestão da informação científica e tecnológica: alguns aportes. **Ciência da Informação**, n. 2, v. 46, 2017.

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, USP, 2008. Disponível em: <http://hdl.handle.net/123456789/809>. Acesso em: 12 out. 2020.

BRASCHER, M.; CARLAN, E. Sistemas de organização do conhecimento: antigas e novas linguagens. In: ROBREDO, J.; BRASCHER, M. (Org.). **Passesios pelos bosques da informação**: estudos sobre representação e organização da informação e do conhecimento. Brasília: IBICT, 2010. p. 147-176. Disponível em: <http://www.ibict.br/publicacoes/eroic.pdf>. Acesso em: 12 out. 2020.

BRASIL. **Decreto** nº 9.812, de 30 de Maio de 2019. Altera o Decreto nº 9.759, de 11 de abril de 2019, que extingue e estabelece diretrizes, regras e limitações para colegiados da administração pública federal. Disponível em [http://www.planalto.gov.br/ccivil\\_03/\\_Ato2019-2022/2019/Decreto/D9812.htm](http://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2019/Decreto/D9812.htm). Acesso em: 15 out. 2020.

BRASIL. **Decreto** nº 10.332, de 28 de Abril de 2020. Institui a Estratégia de Governo Digital para o período de 2020 a 2022, no âmbito dos órgãos e das entidades da administração pública federal direta, autárquica e fundacional e dá outras providências. Disponível em [http://www.planalto.gov.br/ccivil\\_03/\\_Ato2019-2022/2020/Decreto/D10332.htm](http://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2020/Decreto/D10332.htm). Acesso em: 15 out. 2020.

BRASIL. Ministério da Agricultura. Gabinete do Ministro. **Portaria nº 291, de 007 de dezembro de 1990**. Brasília, 1990.

CABRÉ, M. T. **La Terminología**: teoria, metodologia, aplicaciones. Barcelona: Editorial Antártida/Empúries, 1993. 529 p.

CARLAN, E.; MEDEIROS, M. B. B. Sistemas de Organização do Conhecimento na visão da Ciência da Informação. **RICI: R.Ibero-amer. Ci. Inf.**, Brasília, v. 4, n. 2, p. 53-73, ago./dez.2011. Disponível em: [https://repositorio.unb.br/bitstream/10482/12867/1/ARTGO\\_SistemasOrganizacaoConhecimento.pdf](https://repositorio.unb.br/bitstream/10482/12867/1/ARTGO_SistemasOrganizacaoConhecimento.pdf). Acesso em: 10 out. 2020.

CASTRO, R. L.; PELUFE, M. S.; ARRUDA, R. G. Repositórios digitais da Embrapa: acesso livre a produção técnico-científica. In: **CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA, DOCUMENTAÇÃO E CIÊNCIA DA INFORMAÇÃO**, 25, 2013, Florianópolis. Bibliotecas, informação, usuários: abordagens de transformação para a Biblioteconomia e Ciência da Informação: anais. Florianópolis: FEBAB, 2013. Disponível em: <http://ainfo.cnptia.embrapa.br/digital/bitstream/item/86890/1/Repositorios-digitais.pdf>. Acesso em: 15 out. 2020.

CESARINO, M. A. N.; PINTO, M. C. M. F. Cabeçalho de assunto como linguagem de indexação. **Revista da Escola de Biblioteconomia da UFMG**, v. 7, n. 2, 1978. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/71210>. Acesso em: 28 fev. 2021.

CHASTINET, Y. S. et al. Análise da expansão do serviço de bibliografias personalizadas em agricultura (BIP/AGRI): um serviço brasileiro de disseminação seletiva. **Revista de Biblioteconomia de Brasília**, v. 6, n. 2, p. 216-229, 1978.

CINTRA, A. M. M. *et al.* **Para entender as linguagens documentárias**. São Paulo: Editora Polis, 1994.

CINTRA, A. M. M. et al. Linguagem documentária e terminologia. In: **A constituição da normalização terminológica no Brasil**. 2.ed. São Paulo: FFLCH/CITRAT, 2001. p. 17-22.

CORDEIRO, P. P. O papel de uma Biblioteca Nacional no fornecimento de informações aos usuários de uma biblioteca pública: o caso Binagri. Brasília, DF: **SNIDA**, 1981.

DAHLBERG, I. Knowledge Organization: a new science? **Knowl. Org.** v. 33, n. 1, 2006.

EMBRAPA. **Resolução Normativa** nº 184, de 4 de abril de 2019, que Dispõe sobre a Política de Governança de Dados, Informação e Conhecimento da Embrapa, estabelecendo princípios, diretrizes, atribuições e responsabilidades para a gestão de dados, informação e conhecimento, bem como quanto à divulgação de informações relevantes na Empresa. Disponível em: <https://www.embrapa.br/politica-de-governanca-de-dados-informacao-e-conhecimento>. Acesso em: 23 out. 2020.

Escola Nacional de Gestão Agropecuária. **Thesagro**. Disponível em: <http://enagro>.

agricultura.gov.br/glossario/thesagro-pagina. Acesso em: 15 set. 2020.

GARCÍA GUTIÉRREZ, A. L. **Estructura lingüística de la documentación: teoría y método**. Murcia, [Espanha]: Universidade de Murcia, Secretariado de Publicaciones, 1990.

GARDIN, J. Document analysis and linguist theory. **Journal of documentation**, Bradford, [United Kingdon], v. 19, n. 2, p. 137-168, 1973.

GERHARDT, T. E.; SILVEIRA, D. T. **Métodos de Pesquisa**. Porto Alegre: Editora da UFRGS, 2009. 120 p. Disponível em: <http://www.ufrgs.br/cursopgdr/downloadsSerie/derado05.pdf>. Acesso em: 07 set. 2020.

GOMES, H. E. **Classificação, tesauro e terminologia: fundamentos comuns**. Palestra preparada para as Tertúlias do Departamento de Biblioteconomia da UNIRIO, Rio de Janeiro - RJ, 1996. Disponível em: <http://eooci.uff.br/classificacao-tesauro-e-terminologia/>. Acesso em: 15 out. 2020.

MACULAN, B. C. M. S.; DIAS, C. C.; LARA, M. L. G. Subsídios teórico-metodológicos para a construção de vocabulários controlados. In: **Do tratamento à organização da informação: reflexões sobre concepções, perspectivas e tendências**. Coimbra, Portugal: Imprensa da Universidade de Coimbra, 2020. p.193-242.

MARCONI, M. de A.; LAKATOS, E. M. **Fundamentos de metodologia científica**. 7. ed. Editora Atlas, 2010.

MOREIRA, A.; ALVARENGA, L.; OLIVEIRA, A. P. O nível do conhecimento e os instrumentos de representação: tesouros e ontologias. **DataGramZero**, v. 5, n. 6, 2004. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/6563>. Acesso em: 19 out. 2020.

MOREIRA, W. **A construção de informações documentárias: aportes da linguística documentária, da terminologia e das ontologias**. Tese (Doutorado em Ciência da Informação). Universidade de São Paulo, Escola de Comunicação e Artes, 2010.

PINHEIRO, L. V. R.; FERREZ, H. D. **Tesauro Brasileiro de Ciência da Informação**. Rio de Janeiro; Brasília: Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), 2014. Disponível em: [https://ibict.br/images/internas/TESAURO-COMPLETO-FINAL-COM-CAPA-\\_24102014.pdf](https://ibict.br/images/internas/TESAURO-COMPLETO-FINAL-COM-CAPA-_24102014.pdf). Acesso em: 07 set. 2020.

QUIVY, R.; CAMPENHOUDT, L. V. **Manual de Investigação em Ciências Sociais**. Lisboa: Gradiva, 1992.

SALES, R.; CAFÉ, L. Diferenças entre tesouros e ontologias. **Perspect. Ciênc. Inf.**, Belo Horizonte, v. 14, n. 1, pág. 99-116, abril de 2009. Disponível em [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1413-99362009000100008&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-99362009000100008&lng=en&nrm=iso). Acesso: em 10 set. 2020.

UNISIST guidelines for the establishment and development of monolingual thesauri. Paris, **UNESCO**, 1973. 27p.

ZANINELLI, T. B.; CATARINO, M. E. Recuperação de informação no banco de dados da Embrapa Soja: estudo do AINFO. **Informação & Informação**, [S.l.], v. 9, n. 1-2, p. 48-66, jan. 2004. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/1726>. Acesso em: 02 nov. 2020.

ZENG, M. L. Knowledge organization systems (KOS). **KO KNOWLEDGE ORGANIZATION**, v. 35, n. 2-3, p. 160-182, 2008.



# Sistemas de organização do conhecimento: uma análise do Tesouro Unesp

Carlos Víctor de Oliveira<sup>1</sup>, Jamille Abreu Passalini de Sousa<sup>2</sup>  
e Kelly Maria Ayala de Carvalho<sup>3</sup>

---

## 1 Introdução

OS SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO (SOC) SÃO INSTRUMENTOS que dão suporte para a sistematização da organização do conhecimento para recuperação de informação, tanto no ambiente digital como no tradicional. A partir disso, o objetivo deste trabalho é realizar uma análise deste instrumento existente intitulado tesouro da Unesp.<sup>4</sup>

Como metodologia, foi realizado um levantamento bibliográfico sobre organização do conhecimento e sistemas de organização do conhecimento, para contextualização do tema e referencial teórico. Além disso, foi realizada uma análise no portal do Tesouro Unesp observando os critérios de avaliação propostos por Hagar Espanha Gomes, Maria Luiza de Almeida Campos e Dilza Fonseca da Motta (2004) e a aplicação de um questionário elaborado pelos autores à equipe da Unesp que é responsável pelo instrumento.

O capítulo inicia com considerações sobre conceituação de organização do conhecimento segundo teóricos renomados, seguido de considerações breves sobre sistemas de organização do conhecimento (SOC) e adentra o sistema de tesouro. Em seguida, apresenta o estudo de caso do tesouro da Unesp relatando o histórico

---

1 Doutorando em Ciência da Informação - IBICT/UFRJ. Professor do Instituto Federal do Rio de Janeiro. prof.carlosvictor@gmail.com

2 Mestre em Ciência da Informação - IBICT/UFRJ. Arquivista e bibliotecária. j.passalini@gmail.com

3 Mestranda em Ciência da Informação - IBICT/UFRJ. Bibliotecária Digital da Fundação Getulio Vargas. kellyayala10@gmail.com.

4 Agradecemos a ajuda da bibliotecária Rosane Rodrigues de Barros Ribas e da professora Mariângela Fujita que apesar da pandemia do COVID-19 atenderam às nossas solicitações referentes ao Tesouro Unesp.

e informações técnicas e específicas desse soc. A última seção é dedicada às Considerações finais do trabalho.

## **2 Organização do conhecimento**

A organização do conhecimento é definida por Souza (2007) como uma área central de ensino e pesquisa em Ciência da Informação e Biblioteconomia, onde os processos de produção, tratamento e disseminação da informação passaram por mudanças significativas na trajetória da comunicação do conhecimento. Ela identificou quatro fases que marcaram e impulsionaram a criação de instrumentos de classificação e o desenvolvimento de métodos e técnicas de indexação tendo como objetivo a recuperação de documentos e informação no contexto das bibliotecas e de outros sistemas e redes de informação que foram: o ‘Caos’ Documentário, a ‘Explosão’ da Informação, a ‘Avalanche’ de Conhecimento e a ‘Revolução’ Tecnológica.

A relação do ‘caos’ documentário está ligada ao volume crescente de documentos e à diversidade dos tipos de documentos existentes. Quanto à era da ‘explosão’ da Informação, estão presentes a evolução dos métodos e as técnicas de tratamento da informação. A ‘avalanche’ de conhecimento está presente no contexto em que é caracterizada a Ciência da Informação.

Saracevic (1996) conceitua a Ciência da Informação como um campo que busca uma comunicação do conhecimento mais efetiva através da prática profissional e da investigação científica. Essa definição evidencia a importância da Ciência da Informação como suporte profissional e de investigação frente à ‘avalanche’ de conhecimento, e indica a atenção e necessidade devidas à ‘revolução’ tecnológica, cada vez mais marcante a partir da segunda metade do século xx.

Para Souza (2007), a ‘revolução’ tecnológica destaca a ‘Web semântica’ e a ‘Folksonomia’, onde a Web semântica está interessada no refinamento da indexação visando atribuir significado à informação, entendível pela máquina e a Folksonomia é considerada como um novo conceito que está se estabelecendo para significar a organização de recursos digitais na Web. O panorama atual configura então a relevância do papel da organização do conhecimento e da representação da informação na busca por informação no ambiente Web.

Dahlberg (2006, p. 12) enxerga o conhecimento como “a certeza subjetiva e objetiva fundada por alguém sobre a existência de um fato ou de um assunto. Este conhecimento não é transferível, só pode ser elaborado por reflexão pessoal”. Além disso, é necessário definir em qual universo a Organização do Conhecimento se desenvolverá como disciplina, tendo em vista a variedade de aplicações das atividades desenvolvidas por essa área do conhecimento (DAHLBERG, 2006).

A seguir, apresentamos a definição de alguns tipos de SOC que consideramos pertinentes para o desenvolvimento desse trabalho para avançarmos na caracterização do SOC escolhido para análise.

### **3 Sistemas de organização do conhecimento**

A representação da informação se manifesta de diversas formas, o que torna a área e os profissionais de gestão da informação versáteis em criar mecanismos que atendam às necessidades de organização e representação do conhecimento (SOUZA; TUDHOPE; ALMEIDA, 2012). Diante desse cenário, os Sistemas de Organização do Conhecimento são as ferramentas utilizadas nas mais diversas áreas do saber para recuperar, representar e organizar a informação. A função, tipo e estrutura de SOC varia de acordo com o grau de complexidade e especificação da representação.

Assim sendo, os SOC podem ser tesouros, ontologias, taxonomias, vocabulários controlados e outros sistemas de classificação cujo objetivo seja apoiar as atividades de organização e recuperação da informação, a partir do estabelecimento de padrões entre os termos definidos para representar determinados conceitos (MELO; BRASCHER, 2016).

De forma geral, o desenvolvimento dos SOC está relacionado com as perspectivas das abordagens teóricas acerca do conceito e seus relacionamentos, conforme apontam Melo e Brascher (2016). Em função disso, as autoras também afirmam que, apesar das diferenças entre os diversos tipos de SOC, todos possuem uma base conceitual fundamentada em um elemento do conhecimento: o conceito.

Sendo a presente análise especificamente sobre um tesouro, é explorada de maneira breve a Teoria do Conceito, uma vez que se trata de um tema importante para o desenvolvimento de SOC no geral. De acordo com Dahlberg (1978, p. 101), “o conhecimento fixou-se através dos elementos da linguagem” de modo que é necessário fazer uma distinção entre a linguagem empregada no dia a dia e um outro tipo de linguagem, as artificiais ou linguagens formalizadas.

Tendo em vista a necessidade de distinção entre esses dois tipos de linguagem, Dahlberg desenvolve a Teoria do Conceito utilizando o método analítico-sintético para a construção dos conceitos e define conceito como:

Podemos agora definir a formação dos conceitos como a reunião e compilação de enunciados verdadeiros a respeito de determinado objeto. Para fixar o resultado dessa compilação necessitamos de um instrumento. Este é constituído pela palavra ou por qualquer signo que possa traduzir e fixar essa compilação. É possível definir, então, o conceito

como a compilação de enunciados verdadeiros sobre determinado objeto, fixada por um símbolo linguístico (DAHLBERG, 1978, p. 102).

Além dessa definição de conceito, a autora apresenta os elementos constituintes dos conceitos, como os enunciados citados e as características representadas pelos atributos de tais enunciados. Esses elementos são importantes porque indicam as funções das características dos conceitos, bem como a relação entre eles.

A relação entre os conceitos é um tópico importante quando se trata do desenvolvimento de certos SOC, como é o caso dos tesouros. Dentre os relacionamentos citados por Dahlberg (1978) destacamos os seguintes: relação hierárquica (implicação) e relação partitiva. A relação hierárquica entre os conceitos se expressa quando dois conceitos apresentam características iguais, porém um deles apresenta uma característica a mais. Já a relação partitiva se estabelece entre um todo e suas partes, assim como entre um produto e seus elementos constituintes, podendo haver relações hierárquicas, utilizando códigos para o termo genérico (TG) e termo específico (TE) e também relações partitivas entre os termos (lista estruturada). A estrutura dessas relações se configura como uma das maneiras de organização do conhecimento (DAHLBERG, 1978).

#### 4 Os tesouros

A palavra tesouro teve sua origem na Grécia, significando *Treasury or Storehouse* que para o português pode ser traduzido como tesouro ou armazenagem/repositório. Apesar de sua origem, o *Oxford English Dictionary* define em 1936 o termo sendo algo como dicionário ou enciclopédia (DODEBEI, 2014). O termo remete, para alguns, a obra de Roget produzida em 1852 intitulada “*Thesaurus of English Words and Phrases*” que para o próprio autor possuía uma estrutura de classificação verbal, posteriormente discutida sob a luz do pensamento de filósofos como Bacon e Descartes. Esta discussão traçou uma conexão entre Roget e Wilkins, que definiu um novo idioma para substituição do latim a fim de promover o avanço da ciência no século XVII, removendo ruídos dentro da comunicação científica universal (GILCHRIST, 2003).

Os primeiros tesouros não utilizavam técnicas de classificação, seja em sua compilação ou para sua exibição e, apenas com o surgimento do *Thesurofacet* a indexação e recuperação foram claramente demonstradas na combinação de uma classificação facetada e tesouro no mesmo volume, tendo uma correspondência um a um entre facetas e descritores. Inicialmente, utilizavam-se cartões perfurados para a recuperação da informação dentro do classicismo do tesouro, porém com o advento tecnológico o suporte migrou para os computadores (GILCHRIST, 2003).

O tesouro é um SOC que apresenta as relações entre os conceitos de um determinado domínio do conhecimento, sejam essas relações hierárquicas, equivalentes ou associativas. Surge no contexto dos documentos especializados, como resposta à necessidade de recuperação de um grande volume documental (GOMES, 1990).

Dentro dos Sistemas de Organização do Conhecimento, o tesouro é considerado um tipo de vocabulário controlado, que pode ser definido como uma “[...] lista estruturada de conceitos destinados a representar de maneira unívoca o conteúdo dos documentos e das consultas dentro de um sistema documental determinado” (VAN SLYPE, 1991, p. 23-24).

Esta definição também é encontrada na norma da ISO 25964-1, que trata de tesouros e sua interoperabilidade com outros vocabulários, onde o tesouro pode ser definido como “vocabulário controlado e estruturado no qual os conceitos são representados por termos, organizado de forma que as relações entre os conceitos sejam tornadas explícitas e os termos preferenciais sejam acompanhados por entradas iniciais para sinônimos ou quase-sinônimos” (ISO 25964-1, p. 12).

No âmbito da Ciência da Informação, os tesouros ganham notoriedade a partir da década de 1940, conforme afirma Dodebei (2014):

A partir de 1940, o termo tesouro começou a ser utilizado na esfera da Ciência da Informação e, em especial, no processo de recuperação da informação, como sendo um instrumento capaz de transportar conceitos e suas relações mútuas, tal como expressos na linguagem dos documentos, em uma língua regular, com controle de sinônimos e estruturas sintáticas simplificadas (DODEBEI, 2014, p. 66).

A afirmação de Dodebei (2014) demonstra que o propósito do tesouro está relacionado com a recuperação da informação. Para isso, são definidos termos capazes de promover um controle terminológico a partir da tradução de uma “[...] linguagem natural dos documentos, dos indexadores e dos usuários, numa linguagem sistêmica mais rígida” (ROBREDO, 2005, p. 157).

Dessa forma, a construção de um tesouro deve levar em consideração a linguagem do documento, a linguagem dos profissionais da informação e a linguagem dos usuários. Essa característica concede mais uma finalidade ao tesouro, mais propriamente relacionada às atividades dos profissionais da informação. Ou seja, se trata de um SOC que é utilizado tanto pelos usuários com a finalidade de recuperação da informação quanto pelos profissionais da informação para as atividades de indexação (SALES; CAFÉ, 2008).

Ainda sobre as finalidades de um tesouro, é possível apresentar as principais estabelecidas pelo IBICT (1984). De acordo com o documento “Diretrizes para elaboração de tesouros monolíngues” as principais finalidades para a construção de um tesouro são:

- a) controlar os termos usados na indexação mediante um instrumento que traduza a linguagem natural dos autores, usuários e indexadores, para uma linguagem mais controlada; b) uniformizar, mediante esta linguagem documentária, os procedimentos de indexação de profissionais em uma instituição ou numa rede cooperativa; c) limitar o número de termos necessários à explicitação dos conceitos expostos pelos autores de uma área; d) auxiliar a tarefa de recuperação da informação, fornecendo termos adequados para a estratégia de busca (IBICT, 1984, p. 1-2).

O universo conceitual de um tesouro pode se estabelecer de diversas maneiras e é a tarefa mais complexa da elaboração de uma linguagem documentária. Para se determinar este universo podemos obter os termos através da seleção de fontes de informação de dois tipos: primárias e secundárias (SOERGEL, 2014 apud DODEBEI, 2014).

Segundo Dodebei (2014), fontes primárias são caracterizadas pela comunicação do conhecimento em linguagem natural, ao mesmo tempo que as fontes secundárias constituem-se da comunicação do conhecimento em linguagem controlada, ou seja, comunicações condensadas do conhecimento que possuem uma análise estruturada. Se na fonte primária há o ganho de se escolher o termo a partir da literatura e utilizar o termo dos especialistas da área, na secundária há uma preocupação das divergências conceituais existentes entre os conceitos e suas relações por extrair-se o termo de diferentes linguagens documentárias (DODEBEI, 2014).

O princípio da coleta dos termos presentes no universo conceitual para a construção da linguagem documentária possui dois princípios que, idealmente, deveriam existir concomitantemente: a garantia literária e o endosso do usuário. A garantia literária possui como subsídio a literatura, baseando-se a classificação nas classes existentes na literatura e o endosso do usuário é a utilização dos termos presentes na comunidade, onde há uma confirmação direta com os pares da área do conhecimento (DODEBEI, 2014).

Tal combinação apresenta resultados positivos para a combinação de ambos os princípios à medida que, os usuários encontrarão a terminologia preferida utilizada na linguagem documentária e possuirão um maior engajamento no processo e na comunicação com a unidade de informação que elaborou a linguagem documentária (DODEBEI, 2014).

As preocupações a serem consideradas nos mostram a importância de uma metodologia segura para a construção de um tesouro, orientada por normas e padrões amplamente aceitos e difundidos e envolvendo o máximo possível os utilizadores da linguagem, devendo-se sempre promover atualizações e revisões, acompanhando a evolução da área do conhecimento a qual o tesouro pertence.

## **5 O caso Tesouro da Unesp**

A Unesp optou por utilizar o tesouro como um de seus sistemas de organização do conhecimento. Ao acessá-lo, na sua página inicial, a instituição define o tesouro como:

um sistema de organização do conhecimento científico composto de palavras-chave (descritores) relacionadas semântica e genericamente conforme áreas de conhecimento. É utilizado como instrumento para organização, indexação e recuperação da informação em bases de dados. (UNESP, 2019, sem paginação).

O tesouro da instituição foi determinado para trabalhar com termos especializados das áreas de conhecimento de suas atividades de ensino, pesquisa e extensão. Representa o conteúdo da informação mais significativa de livros, dissertações, teses, monografias, trabalhos de conclusão de curso, artigos de periódicos, documentos, legislação etc. que existem na Unesp, e outrossim possui um nível de especificidade e exaustividade necessário para a recuperação da informação de sua comunidade usuária.

O tesouro Unesp é parte fundamental da Política de Indexação das bibliotecas universitárias da Unesp no sentido de orientar seus usuários a como fazer pesquisas e disponibilizar para consulta através da página da web suas palavras-chave. Tais ações têm como objetivo o aprimoramento da recuperação da informação mais precisa por seus usuários durante a busca por assuntos.

## **6 Informações técnicas**

Esta subseção apresenta as informações técnicas do tesouro da Unesp (2021), a partir de sua página da web. O tesouro foi criado em 28 de agosto de 2017 e foi elaborado na língua portuguesa, utilizando como software o TemaTres 3.0 beta. É importante destacar que há possibilidade de comunicação com o tesouro através da API, disponibilizada na plataforma e, apesar do idioma ser o português, é possível visualizar em sua página o tesouro em outros idiomas, sem ser o padrão (português).

Ele se define como uma linguagem do tipo vocabulário controlado, com um total de 89.581 termos e com sua última alteração em 11 de junho de 2019. Colaboraram para a elaboração do Tesouro Unesp ao todo 26 profissionais e 7 bolsistas, somando uma equipe de 33 indivíduos.

Pelo panorama geral oferecido, é possível identificarmos o extenso tamanho do tesouro, com suas relações e termos existentes, sendo um sistema relativamente novo, visto que foi criado em 2017.

## **7 O questionário**

Além da análise no site do tesouro, optou-se também por aplicar um questionário com um colaborador do projeto. O questionário foi respondido por Rosane Rodrigues de Barros Ribas, Bibliotecária, responsável pelo Grupo Técnico de Informação e Documentação da Reitoria, que trabalha na UNESP há 24 anos e no Tesouro desde 2012, quando a Linguagem Unesp começou a ser construída. No cargo de bibliotecária está há 12 anos e como responsável de grupo há 2,5 anos. A seguir a síntese das respostas.

Quando perguntada sobre a quantidade de pessoas que trabalharam na construção do tesouro, respondeu que a construção do tesouro envolveu muitos profissionais entre pesquisadores, bibliotecários, consultores, analistas de sistemas e bolsistas, além dos bibliotecários catalogadores da Rede de Bibliotecas da Unesp e que hoje existe uma Comissão Permanente do Tesouro Unesp com o objetivo de dar continuidade ao trabalho de manutenção e divulgação do tesouro. Porém, a atuação dos bibliotecários da Rede de Bibliotecas da Unesp ainda é fundamental para a manutenção do tesouro.

Rosane disse que a manutenção e atualização se dá por meio de importação de termos da LCSH, MeSH e Biblioteca Nacional, através de correções (compatibilização) em termos já existentes e por meio de criação de termos novos. Quanto à atualização dos termos, ela é feita por bibliotecários e a manutenção do software é feita por analista de sistemas.

Quanto a quantidade de usuários usufruindo do sistema, ela informa que todos os usuários da Rede de Bibliotecas da Unesp usufruem, mesmo que indiretamente do Tesouro Unesp. Ele é condição para a atribuição de termos nos registros bibliográficos dos itens catalogados nas bibliotecas, além de ser fonte de apoio no Repositório Unesp e que o público que utiliza o tesouro é majoritariamente interno, embora ainda não tenha sido feito estudo de uso que está previsto para 2021.

Quando perguntada sobre qual foi o objetivo/motivação para a criação do tesouro, Rosane respondeu que foi por conta da atualização e enriquecimento do vocabulário para melhorar a representação e recuperação.



O motivo que levou a Unesp a decidir trocar o BIBLIODATA pela linguagem UNESP, foi que o BIBLIODATA estava sem atualização estrutural e de vocabulário. Sua configuração e notação não evoluíram conforme sua fonte principal, a LCSH – *Library of Congress Subject Headings*.

Quanto ao software utilizado na criação do tesauro, a resposta foi o software livre TemaTres, porque é de código aberto, permite atualizações, importações e exportações e possui várias possibilidades de formatos diferentes de intercâmbio.

Quando perguntada se foi realizado um estudo de usuário antes da implementação, respondeu que o Tesauro para sua implementação total necessita ainda integrar 5 mil termos ao tesauro Unesp. Porém, já está disponível e em desenvolvimento um estudo de uso do tesauro Unesp cuja aplicabilidade será realizada a partir de 2021.

Para finalizar, foi questionado se as expectativas do tesauro foram alcançadas e Rosane disse que sim, foram alcançadas tendo em vista seu funcionamento junto à ferramenta de interface de busca e seu uso por indexadores da rede de bibliotecas da Unesp e futuramente para autores indexadores no momento do autoarquivamento de suas publicações.

## **8 Análise do site do Tesauro da Unesp**

A análise do site do tesauro da Unesp foi feita utilizando os critérios de avaliação propostos por Hagar Espanha Gomes, Maria Luiza de Almeida Campos e Dilza Fonseca da Motta (2004). A seguir as perguntas utilizadas e as respostas obtidas a partir da análise do site:

### **1. Domínio de Conhecimento coberto pelo tesauro.**

O tesauro da Unesp tem vocabulário com termos especializados das áreas de conhecimento de suas atividades de ensino, pesquisa e extensão.

### **2. Apresenta Introdução? Qual o conteúdo? Está redigida de forma clara a possibilitar o uso do instrumento?**

Possui a seguinte introdução, redigida de forma clara: “É utilizado para representar o conteúdo da informação mais significativa de livros, dissertações, teses, monografias, trabalhos de *conclusão de curso*, *artigos de periódicos*, *documentos*, *legislação e etc com o nível de especificidade e exaustividade necessário para a recuperação da informação de sua comunidade usuária*”

### **3. Forma de apresentação**

#### a) Apresenta parte alfabética

Sim. Em sua página inicial, o tesauro apresenta uma lista dos termos de acordo com a ordem alfabética que, ao ser clicada a letra desejada, exibe uma listagem numerada dos termos relacionados.

#### b) Apresenta parte sistemática

Não há um contexto ou lógica por trás da organização na interface inicial de busca.

### **4. Idioma: monolíngue ou multilíngue?**

Monolíngue.

### **5. Unidade linguística utilizada:**

Assunto

### **6. Quais os tipos de relação encontrados?**

Relação hierárquica (TE e TG), relação associativa (TR) e relação equivalente (USE).

### **7. Aspectos ligados à consistência:**

#### a) Consistência das relações entre os termos?

As relações foram originalmente criadas no bibliodata, e a UNESP começou a desenvolver a partir disso.

É possível notar que Tesauros e Ontologia fazem parte do universo das linguagens documentárias, porém apenas Tesauros possui essa relação hierárquica em seu registro, além disso também possuem inconsistência no plural e singular.

#### b) Consistência no uso do plural e do singular?

Não possui consistência como um todo. Exemplos: Médicos, Cabelo, Olhos, Cerveja, Analgésicos.

Nos exemplos acima, é possível perceber a inconsistência em relação ao plural e singular dos assuntos contemplados pelo tesauro Unesp. Sobre isso, faz-se neces-

sário considerar a literatura da área que dispõe sobre o uso de singular e plural na construção de tesouros, conforme Dodebei (2014).

De certo ponto de vista, nem sempre os termos no plural são os mais gerais, porém para melhor compreendermos esta construção, devemos conhecer a fundo seu significado, pois são pressupostos indispensáveis na construção do domínio perfeito do tesouro. Ou seja, quando não há inconsistência entre os conceitos (DODEBEI, 2014).

### c) Consistência no nível de especificidade?

A própria origem e natureza do tesouro torna essa análise difícil de ser descrita.

#### **8. Nota de aplicação/Esopo: apresenta a definição do termo e/ou a política de indexação?**

Não apresenta a definição do termo e/ou a política de indexação, mas apresenta seus metadados descritivos e permite interoperabilidade com diferentes esquemas.

Podemos constatar que termos específicos, como Cerveja e Cabeçalhos de assunto, não possuem nota de escopo e apenas alguns termos gerais sim, como Administração.

O principal problema identificado não é referente à nota de escopo dos termos específicos, mas da ausência da mesma em alguns termos gerais pois, segundo Dodebei (2014), não é necessária a descrição de todos os termos existentes no tesouro, apenas dos gerais, pois eles devem indicar com clareza quais objetos estão sendo referidos.

#### **9 Considerações finais**

A partir das análises do tesouro da Unesp, podemos ver que o universo conceitual foi oriundo de um universo terminológico já em uso, de fontes secundárias, sem uma participação ativa dos usuários do tesouro, o que para Dodebei (2014) pode ser considerado um ponto chave na elaboração do tesouro junto aos termos presentes na literatura da área.

Tendo em vista as especificidades do Tesouro Unesp apresentadas e o processo de seu desenvolvimento, percebemos que se trata de um soc alinhado com os objetivos da sua instituição mantenedora, uma vez que auxilia a equipe da rede de bibliotecas Unesp como ferramenta para a atividade de indexação, bem como aos usuários do Repositório Institucional na Unesp; motivos pelos quais considera-se que o Tesouro Unesp atendeu as expectativas da equipe responsável por sua implementação.

Da mesma forma, acredita-se que o Tesouro Unesp serve à comunidade de usuários do sistema de bibliotecas, apesar de não estar finalizado. Os próximos pas-

sos envolvem a integração de 5 mil termos e a realização de um estudo de usuários, conforme nos apontou a bibliotecária Rosane Ribas.

O Tesouro da Unesp não atende todos os critérios estabelecidos por Hagar Espanha Gomes, Maria Luiza de Almeida Campos e Dilza Fonseca da Motta (2004) como a apresentação sistemática entre os termos ou consistência no uso do plural e do singular, assim como não apresenta a definição dos termos ou sua política de indexação. Podemos entender isto pelo universo que existe dentro do tesouro, que contempla todas as áreas do conhecimento existentes na universidade.

A utilização de tamanho universo na elaboração de um tesouro acaba por criar alguns empecilhos como a vasta quantidade de termos e diferentes contextos existentes. Portanto, baseado no levantamento bibliográfico realizado, o tesouro Unesp é considerado de fato um tesouro e, sua origem e o tamanho da instituição a qual é vinculado justificam a ausência de critérios, por não se referir a um domínio disciplinar específico, mas a uma instituição que atende a diversas áreas do conhecimento. Essa diversidade torna a tarefa de descrição e análise do tesouro algo difícil de ser explicitado.

### Referências

- 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.
- CAMPOS, M. L. de A., GOMES, H. E., MOTTA, D. F. da. **Elaboração de tesouro documentário: relação entre conceitos e termos.** 2004. Disponível em: [http://www.conexaorio.com/bit/tesouro/criterios\\_aval.htm](http://www.conexaorio.com/bit/tesouro/criterios_aval.htm). Acesso em: 30 set. 2020.
- DAHLBERG, Ingtraut. Teoria do conceito. **Ciência da Informação**, Rio de Janeiro, v.7, n.2, p.101-107, 1978.
- DAHLBERG, I. Knowledge Organization: a new Science? **Knowl. Org.** 33, n.1, 2006 p. 11-19.
- DODEBEI, V. L. D. **Tesouro: linguagem de representação da memória documentária.** Niterói: Intertexto; Rio de Janeiro: Interciência, 2014.
- GILCHRIST, Alan. Thesauri, taxonomies and ontologies – an etymological note. **Journal of Documentation**, v. 59, n. 1, p. 7-18, 2003.
- GOMES, Hagar Espanha. *et al.* **Manual de Elaboração de Tesouros Monolíngues.** Brasília: CNPq/PNBU, 1990. 78p.

IBICT. **Diretrizes para elaboração de tesouros monolíngues**. Brasília, 1984.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO). **ISO 25964 -1: Thesauri for information retrieval**. ISO, 2011.

MELO, M. A. F.; BRÄSCHER, M. **Termo, conceito e relações conceituais**: um estudo das propostas de Dahlberg e Hjørland. *Ciência da Informação*, v. 43, n. 1, 2016.

ROBREDO, J. **Documentação de hoje e de amanhã**: uma abordagem revisitada e contemporânea da Ciência da Informação e de suas aplicações biblioteconômicas, documentárias, arquivistas e museológica. 4. ed. Brasília: Ed. do autor, 2005. 409 p.

SALES, R.; CAFÉ, L. **Semelhanças e diferenças entre tesouros e ontologias**. *DataGramaZero*, v. 9, n. 4, 2008. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/6308>. Acesso em: 30 out. 2020.

SARACEVIC, T. *Ciência da informação: origem, evolução e relações*. **Perspectivas em Ciência da Informação**, v. 1, n. 1, 1996. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/37415>. Acesso em: 1 set. 2020.

SOUZA, Renato Rocha, Douglas Tudhope and Mauricio B. Almeida. 2012. **“Towards a Taxonomy of KOS: Dimensions for Classifying Knowledge Organization Systems”**. *Knowledge Organization* 39, no. 3: 179–192.

SOUZA, R. F. *Organização do Conhecimento*. In: TOUTAIN, L. M. B. B. (org.). **Para entender a ciência da informação**. Salvador: EDUFBA, 2007. p. 103-123.

UNESP. **Tesouro Unesp**. UNESP, 2019. Disponível em: <https://www.biblioteca.unesp.br/tesouro/vocab/index.php>. Acesso em: 20 out. 2020.

VAN SLYPE, G. **Los lenguajes de indización**: concepción, construcción y utilización en los sistemas documentales. Tradução de Pedro Hípola e Félix de Moya. Madri: Fundación Germán Sánchez Ruipérez; Pirâmide, 1991. (Biblioteca del Libro).



# A importância das linguagens documentárias nos ministérios para o acesso à informação no Brasil

Marcelle Costal de Castro dos Santos<sup>1</sup>, Patrícia Conceição Romeu da Fonseca<sup>2</sup>, Priscila Ramos Carvalho<sup>3</sup>

---

## 1 Introdução

A ORGANIZAÇÃO DO CONHECIMENTO E REPRESENTAÇÃO DA INFORMAÇÃO SÃO ATIVIDADES direcionadas para o acesso a documentos (bibliotecas, arquivos, repositórios e sistemas de informação), bem como a recuperação do conteúdo contido nestes documentos (SOUZA, 2007). Neste sentido, os sistemas de classificação possuem um papel importante, pois contribuem para organização, tratamento e recuperação de objetos informacionais físicos e/ou digitais, permitindo o acesso a documentos portadores de informações e conhecimentos.

Diante do “caos documentário” resultante do volume crescente de documentos e da diversidade de tipos de documentos” (SOUZA, 2007, p. 104), a organização do conhecimento tornou-se cada vez mais relevante, visto que adota terminologias padronizadas e o controle da linguagem, eliminando a polissemia e ambiguidades. No contexto da web semântica, a ordenação lógica de acervos de bibliotecas e a organização de informações administrativas de instituições públicas e privadas como, por exemplo, a indexação de informações em bibliotecas digitais e repositórios institucionais, propicia uma melhor representação, recuperação e acesso à informação (SALES; CAMPOS; GOMES, 2008).

Brascher e Café (2010) propõem dois tipos distintos de processos de organização: 1) processo de organização da informação em que se empregam ocorrências individuais de objetos informacionais, e de um conjunto de objetos informacionais para arranjá-los em coleções. Neste caso, temos a organização da informação em bibliotecas, museus, arquivos, tanto tradicionais quanto digitais; 2) processo

---

1 Doutoranda em Ciência da Informação no PPGCI IBICT/ECO-UFRJ, costalcastro@gmail.com

2 Doutoranda em Ciência da Informação no PPGCI IBICT/ECO-UFRJ, romeu.patricia@gmail.com

3 Doutoranda em Ciência da Informação no PPGCI IBICT/ECO-UFRJ, carvalho.priscila@gmail.com

de organização do conhecimento em que se aplica a unidades do pensamento ou conceitos que no que lhe concerne visam à construção de modelos de mundo e constituem-se em abstrações da realidade (BRASCHER; CAFÉ, 2010).

É relevante destacar os problemas terminológicos mencionados por Brascher (1986) na Ciência da Informação, que não diferem dos problemas encontrados nas Ciências Sociais, em relação à imprecisão dos termos e a sua interferência na eficácia da representação e recuperação da informação. Isto ocorre por alguns motivos, tais como: a) uso de termos derivados da linguagem comum que geram diferentes significados para o mesmo termo; b) falta de preocupação com a existência de termos adequados padronizados; c) produtividade terminológica gerando variedade de termos; d) criação de termos novos para representar conceitos já existentes; e) variação semântica do termo resultando na falta de precisão (BRASCHER, 1986).

Em consequência destes desafios terminológicos, Dahlberg (1993) preconiza a teoria do conceito propondo: “qualquer organização do conhecimento deve ser baseada em unidades de conhecimento – que nada mais são do que conceitos” (DAHLBERG, 1993, p. 211)<sup>4</sup>. Para a autora, as unidades de conhecimento consistem em elementos ou características do conceito, os quais contribuem para construção dos sistemas de classificação alinhados como sistemas de conceito. Nessa linha, a escolha de um termo é importante, visto que este proporciona a visibilidade de um conceito em detrimento de outros, o que pode implicar em questões éticas e políticas da escolha, assim como na organização do conhecimento.

No âmbito da administração pública, as dificuldades para recuperar documentos podem resultar em barreiras para o desenvolvimento pleno da democracia, uma vez que podem interferir na transparência das decisões políticas e dos atos administrativos realizados pelo Estado. De acordo com Jardim (1999):

[...] o grau de democratização do Estado encontra, na sua visibilidade, um elemento balizador: maior o acesso à informação governamental, mais democrática as relações entre o Estado e sociedade civil. A visibilidade social do Estado constitui um processo de dimensões políticas, técnicas, tecnológicas e culturais, tendo como um dos seus produtos fundamentais a informação “publicizada” (JARDIM, 1999, p. 49).

---

4 Em inglês: *I would like to add here that the most essential item in the theoretical background of knowledge organization is the fact that any organization of knowledge must be based on knowledge units - which are nothing else but concepts.*



A Lei nº12.527 de 2011, chamada Lei de Acesso à Informação (LAI), foi um marco na transparência dos dados administrativos públicos brasileiros, pois no regime da lei garantiu o acesso à informação de órgãos públicos integrantes dos Poderes Executivo, Legislativo, Judiciário e do Ministério Público. A LAI é fruto do movimento de dados abertos iniciado em 2009, a partir de iniciativas como *My Society* no Reino Unido e *Gov Track* nos Estados Unidos, na linha do Governo 2.0 que disponibiliza o acesso à informação e possibilita o desenvolvimento de novas arquiteturas colaborativas e de participação social.

No cenário da digitalização do conhecimento, a investigação apresenta o resultado de uma pesquisa exploratória e descritiva realizada nos websites dos 17 Ministérios do Brasil, em janeiro de 2021, e pela plataforma de acesso à informação Fala.br, com objetivo de identificar as bibliotecas e as linguagens documentárias, para entender a organização, representação e gestão da informação nos ministérios. Portanto, a seção 2 discorre sobre conceitos que permeiam a organização do conhecimento, compreendendo a complexidade e a importância de cada teoria. A seção 3 expõe os métodos utilizados para mapear os ministérios e, na sequência, a seção 4 exhibe o resultado quantitativo. A seção 5 acrescenta algumas considerações críticas e a última seção mostra as ponderações sobre os resultados encontrados no estudo.

## **2 Abordagens teóricas da representação da informação e do conhecimento**

Desde o século XVI, as classificações bibliográficas têm sido criadas e utilizadas para organização, representação e recuperação da informação em diversos domínios do conhecimento de modo a possibilitar o acesso à informação por usuários como, por exemplo, o tesouro e a tabela de classificação, através de duas formas: alfabética e sistemática (CAMPOS, 1995).

A relação entre termos, palavras e conceitos é apontada por Campos (1995) como primordial na construção de linguagens documentárias. Estas permeiam diferentes perspectivas na Teoria da Classificação Facetada, na Teoria do Conceito e na Teoria Geral da Terminologia, embora as três teorias tenham características semelhantes, “todas objetivam a elaboração de instrumentos de representação do conhecimento” (CAMPOS, 1995, p. 54).

Na Teoria da Classificação Facetada, Ranganathan (1946) sugere para a organização de bibliotecas na Índia o *Colon Classification*, tabela de classificação conhecida por classificação de dois pontos. Para estruturação do conhecimento, Ranganathan dispõe de uma metodologia que considera as características dos conceitos e o relacionamento entre eles como base para a divisão de subclasses (CAMPOS, 1995).

Na Teoria do Conceito, Dahlberg (1978) propõe o uso de elementos de conhecimento para uma série de finalidades diferentes, tais como: análise, construção, reconstrução, comparação, correlação, categorização e definição de conceitos; construção de termos; controle da adequação dos prazos; construção e comparação de sistemas de conceitos. A identificação de elementos do conhecimento e características de conceitos “facilita a compreensão de conceitos em geral, cria uma base para a formação de conceitos e explica a existência de relações entre eles” (DAHLBERG, 1978, p. 150)<sup>5</sup>.

A Teoria Geral da Terminologia (TGT) proposta por Wüster (1998) tem como característica mais relevante a concentração no conceito, sendo o termo a representação do conceito. Nesse sentido, cabe esclarecer os dois objetos da TGT: 1) terminologia como um sistema coeso e sistemático de unidades de linguagem; 2) terminologia concebida como um conjunto de unidades semióticas não necessariamente incluídas na linguagem natural (WÜSTER, 1998).

Cabré (1995) explica a clivagem sobre o objeto de estudo da Lexicologia, na Linguística, trata do estudo das palavras e a Terminologia do estudo dos termos:

Uma palavra é uma unidade descrita por um conjunto de características linguísticas sistemáticas e dotada da propriedade de referir-se a um elemento da realidade. [...] Um termo é uma unidade com características linguísticas semelhantes, usado em um domínio de especialidade (CABRÉ, 1995, p. 7)<sup>6</sup>.

A autora complementa que o campo de trabalho das disciplinas não é coincidente, pois o campo da Lexicologia é mais amplo, e poderia incluir o da Terminologia. Porém, essa lógica só seria possível se as respectivas unidades, o termo e a palavra, fossem o mesmo objeto, opção que Cabré não compartilha, visto que o termo e a palavra são unidades semelhantes, mas, ao mesmo tempo, diferentes (CABRÉ, 1995).

A perspectiva da Terminologia e Lexicologia pode ser diferenciada em razão dos propósitos e objetivos perseguidos por cada campo: a primeira “lida com as

5 Em inglês: *The identification of knowledge elements or characteristics of concepts thus facilitates the understanding of concepts in general, it creates a foundation for the formation of concepts and explains the existence of relationships between them.*

6 Em espanhol: *Una palabra es una unidad descrita por un conjunto de características lingüísticas sistemáticas y dotada de la propiedad de referirse a un elemento de la realidad.[...] Un término es una unidad de características lingüísticas similares, utilizada en un dominio de especialidade.*

palavras para dar conta da competência lexical dos falantes [e a segunda] trata dos termos para definir uma forma de referência” (CABRÉ, 1995, p. 8)<sup>7</sup>.

Convém acrescentar o modelo projetado para demonstrar a natureza multifacetada das relações que existem entre conhecimento, informação e comunicação, no que diz respeito aos aspectos comuns da Terminologia e da Organização do Conhecimento, atuando como elos estruturais de unificação dos três sistemas. Esta abordagem reflete a complexidade e dinâmica inerente na comunicação e nos fluxos de informação, além de demonstrar até que ponto os conceitos científicos, sua representação e sua terminologia técnica fornecem uma estrutura gerenciável para sistemas de conhecimento (BIES, 1997).

Se a Terminologia pode ser definida como a totalidade estruturada dos conceitos presentes em um campo especializado vistos em conjunto com suas respectivas representações (termos e signos não verbais), então é possível considerar que as estruturas terminológicas incorporam a auto-organização inerente aos sistemas. É claro que existe certo perigo escondido em tal declaração, dado que a atribuição de signos e termos a conceitos são uma atividade humana e está sujeita a pontos de vista culturais e individuais específicos. Isto destaca o aumento da complexidade associada aos sistemas de informação multilíngue (BIES, 1997).

É cada vez mais evidente que o problema da representação e da recuperação da informação permeia diversas áreas e extrapola o domínio da Ciência da Informação, o que não significa deixar de lado a aplicação das teorias relacionadas à Organização do Conhecimento. Nos últimos anos, o crescimento da influência das Tecnologias de Informação e Comunicação possibilitou a ampliação do volume e diversidade de formatos de informação, seja pela criação de documentos digitais ou através da digitalização de documentos físicos, propiciando o fenômeno de datificação do conhecimento em escala imensurável e tornando mais complexa a aplicação das três teorias no âmbito científico e técnico-administrativo.

### 3 Metodologia

A pesquisa de natureza empírica teve como objetivo identificar as bibliotecas e as linguagens documentárias ou Sistemas de Organização do Conhecimento (SOC), através de um levantamento de dados nos websites dos 17 Ministérios do Governo Brasileiro e na plataforma de acesso à informação Fala.br, desenvolvida pela Con-

---

7 Em espanhol: *Terminología y lexicología pueden también diferenciarse en razón de los propósitos y objetivos que una y otra persiguen. La lexicología, considerada desde el punto de vista de la lingüística teórica, se ocupa de las palabras con el objetivo de dar cuenta de la competencia léxica de los hablantes; la terminología se ocupa de los términos para fijar una forma de referencia.*

troladoria-Geral da União, que permite cidadãos solicitarem informações públicas de órgãos federais e federados, em conformidade com a Lei de Acesso à Informação e o Código de Defesa dos Usuários de Serviços Públicos.

A partir do prisma da Ciência da Informação, que utiliza técnicas informáticas na aplicação de processos informacionais de representação, organização, gestão e recuperação da informação (ROBREDO; VILAN FILHO, 2010), o estudo utilizou como referencial metodológico a pesquisa exploratória, a fim de “proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito ou a constituir hipóteses” (GIL, 2002, p. 41). Além da pesquisa descritiva que visa a “descrição das características de determinada população ou fenômeno” (GIL, 2002, p. 42).

A coleta de dados foi realizada em janeiro de 2021, tendo como campo de estudo a Organização do Conhecimento nos Ministérios do Governo Brasileiro, com a intenção de compreender a organização, representação e gestão da informação pública nos ministérios através de seus websites. Tomou-se por premissa a existência da biblioteca como centro de informação e conhecimento, bem como a disponibilização de alguma linguagem documentária.

#### 4 Mapeamento das bibliotecas e dos sistemas de organização do conhecimento dos ministérios

O resultado do levantamento nos websites dos 17 ministérios e Fala.br identificou 15 bibliotecas e as linguagens documentárias que se destacaram em quantidade foram: glossário com 19, tesouro com 5, tabela com 1 e dicionário com 1. Além disso, notou-se 7 documentos de vocabulário, os quais servem como suporte na construção das linguagens documentárias. A seguir, o quadro 1 apresenta o resultado consolidado por ministério.

**Quadro 1 - Linguagens Documentárias nos Ministérios Brasileiros.**

Nº	MINISTÉRIO	BIBLIOTECA	LINGUAGEM DOCUMENTÁRIA
1	Ministério Agricultura, Pecuária e Abastecimento (MAPA)	Biblioteca Nacional de Agricultura (Binagri)	Tesouro (Thesagro); Glossário ilustrado de morfologia.
2	Ministério da Cidadania (MC)	Não tem biblioteca	Não possui.
3	Ministério da Ciência, Tecnologia e Inovações (MCTI)	Biblioteca do Ministério da Ciência, Tecnologia e Inovações	Lista de termos controlados.
4	Ministério das Comunicações (MCOM)	Não tem, mas fazem cogestão da Biblioteca do Ministério da Ciência, Tecnologia e Inovações.	Lista de termos controlados.
5	Ministério da Defesa (MD)	Biblioteca do Ministério da Defesa	Glossário das Forças Armadas.

**Quadro 1 - Linguagens Documentárias nos Ministérios Brasileiros.**

Nº	MINISTÉRIO	BIBLIOTECA	LINGUAGEM DOCUMENTÁRIA
6	Ministério do Desenvolvimento Regional (MDR)	Não tem biblioteca física, mas possui biblioteca e videoteca digitais do Programa Nacional de Capacitação das Cidades.	Dicionário do Programa Nacional de Capacitação das Cidades; Glossário do Observatório do Desenvolvimento Regional.
7	Ministério da Economia (ME)	Três bibliotecas físicas: biblioteca Central/DF; biblioteca SRA/RJ; e biblioteca SRA/MG; além de biblioteca digital e Biblioteca Operador Econômico Atualizado (OEA).	Glossário da Receita Federal; Tabela de Atividades; Vocabulário Controlado do Governo Eletrônico (VCGE); Glossário do Patrimônio da União; Sistema de Gestão de Documentos de Arquivo (SIGA).
8	Ministério da Educação (MEC)	Biblioteca física Anísio Teixeira (INEP).	Thesaurus Brasileiro da Educação (Brased); Banco de Dados Terminológicos do Setor Educacional do Mercosul (BDT/SEM); Glossário dos Instrumentos de Avaliação Externa.
9	Ministério de Infraestrutura (MInfra)	Não tem biblioteca física.	Glossário de Termos da Aviação (Anacpédia); Glossário de Termos da Agência Nacional de Transporte Aquaviário (ANTAQ); Glossário de Termos Ferroviários do (Departamento Nacional de Infraestrutura de Transportes DNIT); Glossário de Termos Técnicos Ambientais Rodoviários; Glossário de Termos Técnicos Rodoviários.
10	Ministério da Justiça e Segurança Pública (MJSP)	Biblioteca física integrante da RVBI.	Vocabulário Controlado Básico (VCB).
11	Ministério do Meio Ambiente (MMA)	Biblioteca MMA ou Biblioteca Nacional do Meio Ambiente.	Glossário REDD+; Thesaurus do Meio Ambiente (IBAMA).
12	Ministério de Minas e Energia (MME)	Biblioteca do Ministério de Minas e Energia.	Glossário de Termos Energéticos; Glossário da Agência Nacional de Petróleo (ANP); Vocabulário Controlado da Agência Nacional de Energia Elétrica (ANEEL).
13	Ministério da Mulher, Família e Direitos Humanos. (MMFDH)	Biblioteca do Ministério da Mulher, Família e Direitos Humanos.	Tesouros de Direitos Humanos.
14	Ministério das Relações Exteriores (MRE)	Biblioteca Azevedo da Silveira.	Não foi identificada linguagem documentária.
15	Ministério da Saúde (MS)	Biblioteca física que faz parte da Rede Biblio SUS; e Biblioteca Virtual de Saúde.	Glossário Saúde de A a Z; Tesouro do Ministério da Saúde; Glossários Temáticos; Glossário Eletrônico do Ministério da Saúde; Siglário Eletrônico do Ministério da Saúde.
16	Ministério do Turismo (MTur)	Biblioteca física: Institucional e Demonstrativa do Brasil (SECULT), além de Biblioteca Virtual da Rede de Inteligência de Mercado do Turismo.	Glossário de Turismo.
17	Controladoria Geral da União (CGU)	Biblioteca virtual CGU.	Glossário da CGU.

Fonte: das autoras com base no levantamento de dados (2021).

Nessa lógica, convém explicar o que são as quatro linguagens documentárias mais encontradas nos ministérios: 1) Glossário é uma obra que explica o significado de vocábulos específicos de alguma área, termos técnicos e científicos. É comum aparecer como apêndice de livros técnicos, científicos e obras literárias (CUNHA, CAVALCANTI, 2008); 2) Tesouro é uma linguagem documentária de especialidade do objeto/referente, evidenciando a relação de termos de um domínio entre si, para indexação e recuperação em um sistema de informação (CAMPOS; GOMES, 2006); 3) Tabela é uma representação de dados, números ou conhecimentos dispostos em forma de lista, com colunas paralelas ou linhas horizontais, utilizada para complementar um esquema de classificação. Desta maneira, ampliando a notação de assuntos específicos, como as tabelas auxiliares da Classificação Decimal Universal (CUNHA, CAVALCANTI, 2008); 4) Dicionário é uma obra de referência que traz informações sobre as palavras, sua grafia, pronúncia, significado, etimologia, sinonímia e antonímia, em geral, dispostos na ordem alfabética. Teoricamente, o dicionário analisa apenas a “palavra”, não se ocupando da “coisa”, mas os dicionários modernos, na maioria das vezes, trazem informações sobre a coisa ou objeto, além da palavra (CUNHA, CAVALCANTI, 2008, p. 123).

## 5 Discussão dos resultados

O levantamento de dados nos ministérios possibilitou a percepção da questão do acesso à informação e transparência na gestão da informação pelo Estado. Muitas vezes foi um desafio identificar as bibliotecas e as linguagens documentárias nos websites e, por este motivo, foi utilizado o sistema Fala.br. A seguir apresentamos algumas considerações de cada ministério.

O MAPA respondeu à solicitação no Fala.br em 08 de fevereiro de 2021, na qual confirmou a existência da Biblioteca Nacional de Agricultura, com 112 anos, e o Thesagro, criado por ela em 2016, sendo este uma evolução do Thesaurus de Indexação/Recuperação da Literatura Agrícola Brasileira, de 1979. Foi possível identificar também no website o Glossário ilustrado de Morfologia vinculado à Secretaria de Defesa Agropecuária. É factível inferir que, aparentemente, o ministério demonstra interesse no uso de linguagens documentárias como meio de contribuir para o desenvolvimento do setor.

O MC respondeu em 11 de fevereiro 2021 sinalizando não possuir biblioteca e, por isso, não usa linguagens documentárias. A investigação mostrou que após a fusão do Ministério do Esporte com o Ministério do Desenvolvimento Social, as bibliotecas foram dissolvidas. Infelizmente não foi possível identificar se existia alguma iniciativa de desenvolvimento de linguagem documentária nas antigas bibliotecas.

O MCTI respondeu em 3 de fevereiro de 2021 confirmando a existência da biblioteca, do repositório digital e o uso de lista de termos controlados através do software de gestão de biblioteca SophiA. A pesquisa apontou que devido à diversidade de instituições associadas e voltadas para pesquisa e desenvolvimento científico do país, as linguagens documentárias no ministério podem ser mais exploradas.

O MCOM respondeu à solicitação em 22 de fevereiro de 2021 e discorreu sobre a responsabilidade de gestão do acervo bibliográfico da Divisão de Gestão da Informação, que utiliza o sistema SophiA. No entanto, devido à sua recriação, a biblioteca funciona com regime de cogestão entre MCOM e MCTI. O estudo compreendeu que os sistemas de gestão bibliográficos, biblioteca e repositório digital são compartilhados em decorrência de reestruturações ministeriais. Em virtude da enorme demanda institucional de recursos para desmembrar os sistemas de gestão, acredita-se ser uma medida acertada manter a partilha atual.

O MD enviou a resposta em 10 de fevereiro de 2021, informando apenas o uso do sistema Pergamum da Pontifícia Universidade Católica (PUC) do Paraná. Porém, a pesquisa identificou no website a Rede de Bibliotecas do Ministério da Defesa, criado em 2011, com intuito de integrar os acervos bibliográficos de 91 bibliotecas das três forças armadas, além do glossário. É possível inferir que existam diversas linguagens documentárias não informadas devido à gama de bibliotecas que compõem a rede, sendo um campo vasto a ser explorado.

O MDR respondeu ao pedido em 26 de janeiro de 2021, e comunicou o fechamento da biblioteca em maio de 2020, por falta de bibliotecário. Através do website foi possível notar uma biblioteca, uma videoteca e um dicionário digital no Programa Nacional de Capacitação de Cidades, bem como o glossário do Observatório de Desenvolvimento Regional de 2014, sendo que este soc contempla mais itens. Em razão de a pasta abarcar uma diversidade de temas de outros ministérios, a escolha do glossário para cada tema é acertada.

O ME atendeu ao pedido em 26 de janeiro de 2021, e indicou um link sobre bibliotecas no website, bem como a não utilização de linguagens documentárias. Foi possível constatar a existência de três bibliotecas físicas, sendo duas advindas do extinto Ministério da Fazenda (Rio de Janeiro e Minas Gerais), e uma biblioteca central no Distrito Federal. Além disso, também foram identificados uma biblioteca e um repositório digitais, criados em 2019. Embora não tenham informado o uso de linguagens documentárias, foram encontrados dois glossários (Receita Federal e Patrimônio da União). Ficou clara a dificuldade de se encontrar informações no website e a complexidade para a padronização por linguagens documentárias, devido à incorporação de diversas pastas pelo ministério.

O MEC informou em 26 de janeiro de 2021 o encaminhamento da demanda para o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), o qual prontamente explicou em 19 de fevereiro de 2021 que não existe biblioteca no MEC, mas sim no INEP, bem como o Thesaurus Brasileiro da Educação (Brased) e o Banco de Dados Terminológicos do Setor Educacional do Mercosul (BDT/SEM). O tesouro é composto por termos e conceitos provenientes do Centro de Informação e Biblioteca em Educação (Cibec), sendo possível seu acesso através do Pergamum. É interessante perceber que as informações relacionadas às bibliotecas e linguagens documentárias não estão disponíveis diretamente no website do MEC, dificultando assim o acesso à informação e exigindo de certa forma uma competência em informação por parte dos cidadãos.

O MInfra respondeu à solicitação em 10 de fevereiro de 2021 e informou o fechamento da biblioteca em outubro de 2019. Por conta da absorção do antigo Ministério dos Transportes, a pasta incorporou suas linguagens documentárias, tendo como destaque os glossários especializados de transporte (aviação, aquaviário, ferroviário e rodoviário). Apesar da adequação da linguagem documentária, é plausível sugerir o uso de sistemas que abarquem os termos e conceitos de transporte de maneira interativa, como uso de softwares próprios para o desenvolvimento de tesouros.

O MJSP discorreu em 11 de fevereiro de 2021 sobre a biblioteca criada na gestão do presidente Getúlio Vargas em 1940, que é integrada à Rede Virtual de Bibliotecas (RVBI) e coordenada pelo Senado Federal. Foi constatado também o uso do Vocabulário Controlado Básico (VCB), que visa à uniformidade da indexação e recuperação da informação, composto por descritores da área Jurídica, Ciências Sociais e Humanas. O ministério disponibiliza um documento, acessível online, explicando o seu desenvolvimento e demonstrando certo avanço na organização do conhecimento em comparação a outros ministérios.

O MMA não disponibilizou informações sobre a biblioteca no novo padrão de website, mas no antigo foi possível identificar a biblioteca digital aberta. Contudo, não ficou evidente a linguagem documentária utilizada. Após a resposta do pedido no Fala.br, o ministério sinalizou que sua biblioteca foi transferida para a biblioteca do Instituto Brasileiro do Meio Ambiente e dos Recursos Renováveis (IBAMA). No website do instituto foi possível perceber a biblioteca e o Tesouro do Meio Ambiente com ordenação sistemática e alfabética, mas aparentemente precisa de atualização de termos como, por exemplo, “mudança climática”. Foi encontrado também o Glossário REED+ Brasil, que teve sua última atualização em 2019, em que se notou a falta de padronização de termos indexados.

O MME respondeu à consulta em 11 de fevereiro de 2021 e indicou uma biblioteca e o uso do tesouro da Biblioteca do Senado Federal. No website do ministé-



rio foi encontrado alguns soc em órgãos vinculados através do sistema de busca, tais como: Glossário de Termos Energéticos (atualizado pela última vez em 2003), Glossário de Termo da Agência Nacional de Petróleo (atualizado pela última vez em 2020), e Vocabulário controlado da Agência Nacional de Energia Elétrica. Vale sinalizar a dificuldade de encontrar informações no website sobre a biblioteca e os soc. É admissível sugerir a integração dos setores de Petróleo e Energia, por serem transversais e relevantes para o Estado, por meio de uma taxonomia para o ministério.

O MFDH comunicou em 7 de fevereiro de 2021 que possui uma biblioteca e utiliza um tesouro. O Tesouro de Direitos Humanos é desenvolvido no software TemaTres, integrado ao Koha e Dspace, com o objetivo de padronizar os termos no escopo da atuação da biblioteca. É factível mencionar o desafio da criação de um tesouro que englobe um tema tão amplo.

O MRE possui a Biblioteca Azeredo da Silveira, mas na página da biblioteca não foi identificado o uso de soc. No catálogo da biblioteca, que utiliza o software Pergamum, os termos são indexados em um catálogo de assuntos, visível para o usuário. A partir de busca empírica por termos foi possível perceber duplicações, sinonímia e erros de grafia, o que indica a necessidade de atualização.

O MS respondeu à solicitação em 26 de fevereiro de 2021 e comunicou que a biblioteca do ministério usa o gerenciador abcd, software formado pelas bases de dados Lidbi Web da Rede de Bibliotecas Virtuais em Saúde, e o vocabulário controlado dos Descritores em Ciências da Saúde (DeCS). No website foi encontrado o Glossário de Práticas Integrativas e Complementares da Saúde de 2018, parte de uma série de glossários temáticos no âmbito do Projeto de Terminologias da Saúde, realizado por meio da Coordenação de Disseminação de Informações Técnico-Científicas em Saúde (CODINF). Ademais, no website antigo do MS foi identificada a Rede BiblioSUS. Vale sinalizar que o e-Glossário do ministério parece estar descontinuado, pois os verbetes não estão disponíveis. Já o Siglário disponibiliza cerca de duas mil siglas utilizadas de forma simples, com índice de siglas e palavras. É razoável dizer que o campo da saúde tem avançado no uso de linguagens documentárias, mas no website as informações não estão compreensíveis e atualizadas, talvez devido à migração da plataforma antiga para nova.

O MTur esclareceu em 7 de fevereiro de 2021 que possui a Biblioteca Institucional e a Biblioteca Demonstrativa Maria da Conceição Moreira Salles, uma das mais antigas de Brasília, vinculadas à Secretaria de Estado de Cultura (SECULT). Além disso, não utilizam soc e aguardam a implantação do software Koha para gestão das bibliotecas. No website foram identificados a Biblioteca Virtual da Rede de Inteligência de Mercado do Turismo e o Glossário de Turismo, com termos téc-

nicos em ordem alfabética, definições e referências. Foi possível notar um aparente interesse na padronização de termos do setor visando uma melhor comunicação, embora ainda existam informações no website antigo que não foram migradas para novo, o que dificulta o acesso à informação.

A CGU em resposta ao pedido do Fala.br comunicou a existência do repositório de conhecimento institucional hospedado no Dspace e os documentos são acessados por busca, linha do tempo, tipo de assunto, área temática, local da edição e ano de publicação. O repositório é uma ferramenta para a gestão do conhecimento da pasta, e é aceitável ponderar que a CGU, por ser um órgão de controladoria, tem buscado utilizar os Sistemas de Organização do Conhecimento.

## 6 Considerações finais

A pesquisa percebeu a redução do número de bibliotecas devido à mudança de organização dos ministérios conforme o governo vigente (2019-2022), o que pode ter implicado no resultado do número de linguagens documentárias. Além disso, foi um desafio identificar as bibliotecas e as linguagens documentárias mesmo com a padronização visual de informações nos websites governamentais e a resposta do Fala.br por todos os ministérios.

Em relação à organização, representação e gestão da informação pública no Estado, o estudo aponta a necessidade da ampliação do uso de linguagens documentárias nos ministérios de maneira a permitir uma representação, recuperação e acesso à informação mais eficiente para os servidores públicos e para os cidadãos, assim contribuindo com a Lei de Acesso à Informação.

## Referências

BIES, W. Book reviews. **Knowledge Organization**, [S. l], v. 24, n. 3, p.190-193, 1997. DOI: <https://doi.org/10.5771/0943-7444-1997>

BRASCHER, M. Terminologia brasileira em Ciência da Informação: uma análise. **Ciência Da Informação**, Brasília, DF, v. 15, n. 2, p. 135-142, 1986.

BRASCHER, M.; CAFÉ, L. Organização da Informação ou Organização do Conhecimento? In: LARA M. L. G; SMIT, J. W. (org.). **Temas de pesquisa em Ciência da Informação no Brasil**. São Paulo: Escola de Comunicações e Artes da USP, 2010. p. 85-103.

BRASIL. **Lei nº 12.527, 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; altera a Lei nº 8.112, de 11 de dezembro de 1990;

revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991; e dá outras providências. Brasília, DF: Presidência da República, 2011. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2011/lei/l12527.htm](http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm). Acesso em: 9 set. 2021.

CABRÉ, M. T. La terminología hoy: concepciones, tendencias y aplicaciones. **Ciência Da Informação**, Brasília, DF, v. 24, n. 3, [p. 1-15], 1995.

CAMPOS, M. L. A. Linguagens documentárias: núcleo básico de conhecimento para seu estudo. **Revista da Escola de Biblioteconomia da UFMG**, Belo Horizonte, v. 24, n. 1, p. 52-62, 1995. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/76251>. Acesso em: 26 ago. 2021.

CAMPOS, M. L. A.; GOMES, H. E. Metodologia de elaboração de tesauro conceitual: a categorização como princípio norteador. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 11, n. 3, p. 348-359, 2006.

CUNHA, M. B.; CAVALCANTI, C. R. de O. **Dicionário de Biblioteconomia e Arquivologia**. Brasília, DF: Briquet de Lemos, 2008.

DAHLBERG, I. Knowledge organization: its scope and possibilities. **Knowledge Organization**, [S. l], v. 20, n. 4, p. 211-222, 1993.

DAHLBERG, I. A referent-oriented analytical concept theory of Interconcept. **International Classification**, [S. l], v. 5, n. 3, p. 142-150, 1978.

GIL, A. C. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2002.

JARDIM, J. M. **Transparência e opacidade do Estado no Brasil**: usos e desusos da informação governamental. Niterói: Editora da Universidade Federal Fluminense, 1999.

RANGANATHAN, S. R. **Suggestions for the organization of libraries**. India: Oxford University Press, 1946.

ROBREDO, J.; VILAN FILHO, J. L. Metrias da informação: história e tendências. In: ROBREDO, J.; BRÄSCHER, M. (org.). **Passeios pelo bosque da informação**: estudos sobre a representação e organização da informação e do conhecimento. Brasília, DF: IBICT, 2010. p. 1-335.

SALES, L. F.; CAMPOS, M. L. A.; GOMES, H. E. Ontologias de domínio: um estudo das relações conceituais. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 13, n. 2, p. 62-76, 2008. DOI: <https://doi.org/10.1590/>

S1413-99362008000200006. Disponível em: <https://www.scielo.br/j/pci/a/qNXG6bfyKvsmtmKTcn6MwCp/?lang=pt#>. Acesso em: 9 set. 2021.

SOUZA, R. F. de. Organização do conhecimento. *In*: TOUTAIN, L. M. B. B. (org.). **Para entender a Ciência da Informação**. Salvador: EDUFBA, 2007. p. 103-124.

WÜSTER, E. **Introducción a la teoría general de la terminología y a la lexicografía terminológica**. Barcelona: IULA, 1998.



50

---

## Realização

---



---

## Cooperação



Cooperação  
Representação  
no Brasil



---

## Financiamento

---

---

ESTA OBRA É PARTE DA COLEÇÃO PPGCI 50 ANOS E FOI COMPOSTA EM MINION PELO PROGRAMA DE EDUCAÇÃO TUTORIAL DA ESCOLA DE COMUNICAÇÃO DA UFRJ EM SETEMBRO DE 2021.



Este livro é o produto da experiência de ensino da disciplina Sistemas de Organização do Conhecimento ministrada pelos professores Rosali Fernandez de Souza, Gustavo Saldanha e Luana Sales em 2020. Os artigos apresentam os relatos das pesquisas desenvolvidas pelos alunos sobre o uso de SOC em ambientes acadêmicos e em órgãos da administração pública federal. Através deles, são reveladas diferentes abordagens metodológicas de coleta e análise dos dados adotadas pelos grupos de trabalho, cujos resultados e discussões podem não apenas indicar a relevância do desempenho de SOC nos espaços estudados, mas também motivar pesquisas em outros ambientes de informação.

EM COOPERAÇÃO

