

RODRIGO DE SANTIS

SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO PARA
DOMÍNIOS COMPLEXOS: abordagem a canções populares na web
semântica utilizando propriedades fuzzy

Tese de doutorado
Julho de 2016



UNIVERSIDADE FEDERAL DO RIO DE JANEIRO – UFRJ
ESCOLA DE COMUNICAÇÃO - ECO
INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA - IBICT
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO - PPGCI

RODRIGO DE SANTIS

SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO PARA DOMÍNIOS
COMPLEXOS: abordagem a canções populares na web semântica utilizando propriedades
fuzzy

RIO DE JANEIRO

2016

S236 De Santis, Rodrigo.
Sistemas de organização do conhecimento para domínios complexos: abordagem a canções populares na web semântica utilizando propriedades fuzzy / Rodrigo De Santis. 2016.
149f. : il..

Orientadora: Rosali Fernandez de Souza.

Tese (doutorado) – Universidade Federal do Rio de Janeiro, Escola de Comunicação; Instituto Brasileiro de Informação em Ciência e Tecnologia, Programa de Pós Graduação em Ciência da Informação, 2016.

1. Ciência da informação. 2. Organização do conhecimento. 3. Sistemas de Informação. 4. Classificação 5. Semiótica. I. Souza, Rosali Fernandez de. III. Universidade Federal do Rio de Janeiro. Escola de Comunicação. II. IBICT.

CDD: 020.1

RODRIGO DE SANTIS

**SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO PARA DOMÍNIOS
COMPLEXOS: abordagem a canções populares na web semântica utilizando propriedades
fuzzy**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciência da Informação, convênio entre o Instituto Brasileiro de Informação em Ciência e Tecnologia e a Universidade Federal do Rio de Janeiro / Escola de Comunicação, como requisito parcial à obtenção do título de Doutor em Ciência da Informação.

Orientadora: Rosali Fernandez de Souza

**RIO DE JANEIRO
2016**

RODRIGO DE SANTIS

SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO PARA DOMÍNIOS

**COMPLEXOS: abordagem a canções populares na web semântica utilizando propriedades
fuzzy**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciência da Informação, convênio entre o Instituto Brasileiro de Informação em Ciência e Tecnologia e a Universidade Federal do Rio de Janeiro / Escola de Comunicação, como requisito parcial à obtenção do título de Doutor em Ciência da Informação.

Rio de Janeiro, 7 de julho de 2016

Profa. Dra. Rosali Fernandez de Souza, orientadora
PPGCI/IBICT – ECO/UFRJ

Profa. Dra. Rosana Saldanha da Gama Lanzelotte
PPGM/UNIRIO – Instituto Musica Brasilis

Profa. Dra. Vera Lucia Doyle Louzada de Mattos Dodebei
PPGMS/UNIRIO

Profa. Dra. Maria Nélida González de Gómez
PPGCI/IBICT – ECO/UFRJ – PPGCI/UFF

Profa. Dra. Regina Maria Marteleto
PPGCI/IBICT – ECO/UFRJ

Prof. Dr. Gustavo Silva Saldanha
PPGCI/IBICT – ECO/UFRJ

Para Adrianna, companheira na vida,
no amor e nas canções

AGRADECIMENTOS

Escrever uma tese de doutorado é um trabalho individual, solitário e que, sem nenhuma contradição, só é possível com a participação e o apoio de muitas pessoas e instituições. A seguir, deixo registrados meus sinceros agradecimentos a algumas delas.

Início agradecendo à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela bolsa a mim concedida durante 18 meses, período no qual pude participar de eventos científicos, adquirir material bibliográfico e dedicar-me de maneira ainda mais concentrada à pesquisa.

Ainda do ponto de vista institucional, agradeço ao IBICT e à UFRJ, mas na pessoa de seus servidores (professores, secretários, coordenadores, bibliotecários, faxineiros...) – todos elementos fundamentais para a existência e funcionamento destas instituições e para a consecução de sua missão de oferecer cursos de elevada qualidade, como este em que tive o privilégio de ser aluno.

Agradeço aos colegas de curso pelo incentivo, pelas conversas e por todas as experiências trocadas. Agradeço a todos e peço que se sintam representados pelos que cito nominalmente: Thulio Gomes, Rodrigo Bozzetti, Heloisa Ottoni, Moisés Nisenbaum, Carlos Vieira e Vinícios Menezes.

A seguir, agradeço pelo privilégio de poder conviver com pessoas possuidoras de virtudes raras, capazes de combinar medidas igualmente grandes de competência e de generosidade. A cada uma delas, a minha mais profunda gratidão.

Pela minha trajetória de estudos na Informática, na Música e na Semiótica, manifesto meus agradecimentos aos professores que ajudaram a construir as bases desta pesquisa, especialmente Sean Siqueira, Angelo Ciarlini, Carlos Alberto Figueiredo, Lucia Teixeira, Renata Mancini e José Roberto do Carmo Jr.

Agradeço aos professores José Augusto Chaves Guimarães e Vera Dodebei pela acolhida calorosa que tive desde o primeiro evento em que participei na área de Organização do Conhecimento - o congresso do capítulo brasileiro da ISKO em 2013. À Vera eu agradeço desde antes, quando terminado o meu mestrado na UNIRIO, abriu as portas das Ciências Sociais Aplicadas ao me receber em reuniões de seu grupo na Memória Social.

Também agradeço, desde a chegada, à professora Maria Nélida González de Gómez que, naquele mesmo congresso da ISKO, teve a generosidade de comentar a minha apresentação, tecer considerações sobre a aproximação que eu propunha entre a Organização do Conhecimento e a noção de formação discursiva em Foucault. Seus comentários e a leitura de seus textos me incentivaram a continuar a pesquisa e a estreitar o diálogo com o domínio filosófico.

Agradeço também a dois pesquisadores italianos da Organização do Conhecimento, receptivos às minhas ideias e abertos a compartilhar suas pesquisas: Fulvio Mazzocchi e Claudio Gnoli. *Grazie!*

Agradeço a todos os professores do IBICT pelo diálogo estabelecido, pelas disciplinas oferecidas e pela confiança em mim depositada.

Agradeço nominalmente aos professores Marco Schneider e Gustavo Saldanha, cujo curso “Teoria e História da Informação” foi crucial na minha formação e me fez enxergar os estudos sobre informação a partir de um olhar mais crítico e mais abrangente. Ao Gustavo quero agradecer de forma ainda mais especial, pela enorme disponibilidade e invencível disposição em ajudar, ler, ouvir, comentar, participar, compartilhar, sugerir, corrigir e orientar.

Agradeço a Regina Marteleto, professora na mais profunda acepção do termo, pela capacidade de formação e transformação. Lembro-me do receio que tive ao me inscrever em seu curso – como eu, um neófito na área poderia acompanhar leituras tão difíceis como as de Pierre Bourdieu, Walter Benjamin, Georg Simmel propostas em sua disciplina? Regina (“não me chamem de senhora”) desde o primeiro contato desmistificou o meu temor. Sinto-me verdadeiramente realizado por poder utilizar nesta tese parte significativa do que aprendi em suas aulas. Obrigado Regina pelas aulas, pelas leituras sugeridas, pelos seminários e pela oportunidade de participar em seu grupo de pesquisa. Em suma, obrigado pelo convívio profícuo e amável.

De forma a destacar ainda mais sua importância, deixo para o final os agradecimentos àquelas pessoas sem as quais, sem nenhum exagero, eu não teria conseguido chegar até aqui, que são insubstituíveis na minha vida e na minha trajetória.

Em primeiro lugar, agradeço à minha orientadora, Rosali, profissional competente e pessoa de insuperável gentileza em todos os momentos. Agradeço pela receptividade na Organização do Conhecimento, pela paciência em me apresentar e conduzir na caminhada pela nova estrada e pela sempre agradável e profícua relação de orientação que se transformou em sincera amizade, plena de respeito, admiração e gratidão. Prova que, assim como a ideia defendida nesta tese, as fronteiras das relações humanas são fuzzy, e os domínios acadêmicos, da amizade e da família (incluindo o Pitufo, claro) se inter-relacionam de maneira a constituir um todo harmônico e frutífero.

Agradeço a Rosana, minha orientadora “a partir” do mestrado, com quem tenho a honra e a alegria de trabalhar desde então. À Rosana devo a base dos méritos científicos e acadêmicos que porventura tenha alcançado. Sua generosidade e a disposição em compartilhar sua imensurável competência e capacidade intelectual, cultural e humana me formaram como pesquisador e me influenciam como indivíduo no mundo. Seu entusiasmo pela pesquisa, pela ciência, pela música, e a sua coragem de abraçar projetos desafiadores que são “para a vida toda”, me inspiram a seguir pelo caminho do conhecimento. Agradeço pelos quase dez anos de parceria, que espero poder manter “para a vida toda”.

Para a vida toda agradeço também aos familiares e amigos que participaram ao longo destes anos para a realização desta pesquisa. E para que todos possam se sentir devidamente abraçados, escolho agradecer nominalmente à vovó “dona” Glória, por tudo o que significa, pela ligação com o sagrado e pelas bases familiares e de amizade que perpassam gerações.

DE SANTIS, Rodrigo. *Sistemas de organização do conhecimento para domínios complexos: abordagem a canções populares na web semântica utilizando propriedades fuzzy*. Orientadora: Rosali Fernandez de Souza. Tese (Doutorado em Ciência da Informação) – Programa de Pós Graduação em Ciência da Informação – Escola de Comunicação / Universidade Federal do Rio de Janeiro – Instituto Brasileiro de Informação em Ciência e Tecnologia, Rio de Janeiro, 2016.

RESUMO

A ciência contemporânea vem se deparando com a necessidade de novas formulações que sejam capazes de superar as restrições decorrentes do pensamento científico tradicional, de inspiração cartesiana e que postula: a separação clara e inequívoca dos objetos do mundo; a possibilidade de divisão do todo em partes menores; a resolução de problemas a partir das partes menores; a garantia de recomposição da totalidade a partir da junção das partes. Teorias do pensamento complexo vem sendo desenvolvidas como alternativas a este modelo simplificador, no intuito de evidenciar as suas limitações e oferecer métodos de abordagem compatíveis com este novo paradigma. Esta pesquisa investiga as bases para concepção de um sistema de organização do conhecimento capaz de lidar com domínios complexos, identificando as matrizes de sua formulação em autores como conde de Buffon e Ranganathan. A superação das dicotomias e da noção teleológica do conhecimento é investigada a partir de uma abordagem semiótica, fundamentada na representação simbólica do conhecimento, a partir dos trabalhos de Ernst Cassirer e baseia-se na noção de fronteira em Gilles Deleuze e de dispositivo em Michel Foucault. O método complexo proposto por Edgar Morin fundamenta a base metodológica para aplicação do esquematismo tensivo proposto por Viggo Brøndal. A formulação de um sistema de organização do conhecimento capaz de abordar domínios complexos baseia-se em um duplo deslocamento: o primeiro consiste no reconhecimento da primazia da recuperação em relação à representação em sistemas de organização do conhecimento, hipótese defendida desde a década de 1950 por pesquisadores como Jesse Shera e posteriormente por Brian Vickery. O segundo deslocamento diz respeito à unidade básica do conhecimento tratável, que deixa de ser a noção de conceito – hierárquico, dicotômico, com fronteiras bem definidas e cujo pertencimento é reduzido a uma relação binária, de valor verdadeiro ou falso –, para se tornar o dispositivo infocomunicacional – com suas relações imanentes, potencialmente desdobráveis ao infinito, com fronteiras sobrepostas e mutuamente interferentes, e o pertencimento mensurável em graus. O recurso a ferramentas e linguagens matemáticas e tecnológicas, como a lógica fuzzy e a web semântica, possibilitaram a consecução prática deste tipo de abordagem, que foi aplicada empiricamente no domínio da canção popular brasileira. Este domínio complexo, dotado de uma linguagem própria (a linguagem da canção) permitiu verificar a insuficiência das abordagens tradicionais, baseadas em metadados ou métodos colaborativos de classificação e mostrou-se adequado ao permitir uma abordagem multidimensional do conhecimento, considerando-o não como produto encerrado em si, mas enfatizando seu caráter dinâmico enquanto processo cuja construção é de natureza social e histórica.

Palavras-chave: Dispositivo infocomunicacional. Rizoma. Semiótica. Classificação. OWL.

DE SANTIS, Rodrigo. *Knowledge organization systems for complex domains: an approach to popular songs in the semantic web using fuzzy properties*. Supervisor: Rosali Fernandez de Souza. Thesis (PhD in Information Science) – Information Science Post Graduation Program – Communication School / Federal University of Rio de Janeiro – Brazilian Institute of Information in Science and Technology, Rio de Janeiro, 2016.

ABSTRACT

Contemporary science has been faced with the need of new approaches, able to overcome the constraints resulting from the traditional scientific thought of cartesian inspiration and that postulates: the clear and unambiguous separation of objects in the world; the possibility of division of the whole into smaller parts; the problem solving from the smaller parts; the guarantee of full recovery of the totality from the connection of the parties. Theories of complex thinking has been developed as alternatives to this simplifying model in order to show its limitations and offer approach methods compatible with this new paradigm. This research investigates the basis for the conception of an knowledge organization systems capable of dealing with complex domains, identifying the foundation of its formulation in authors as comte de Buffon and Ranganathan. The overcoming of dichotomies and of the teleological notion of knowledge is investigated from a semiotic approach, based on symbolic representation of knowledge, from Ernst Cassirer works, and based on the use of the notion of boundaries, by Gilles Deleuze, and of dispositif, by Michel Foucault. The complex method proposed by Edgar Morin is the methodological approach for the application of tensive schematism proposed by Viggo Brøndal. The formulation of an organization knowledge system able to address complex domains is based on a double displacement: the first one consist in the recognition of the primacy of recovery in relation to the representation in organizational systems of knowledge, hypothesis advocated since the 1950s by researchers as Jesse Shera and later by Brian Vickery. The second displacement concerns to the basic unit of treatable knowledge, that is no longer the notion of concept – hierarchical, dichotomous, with its well defined boundaries and which belonging is reduced to a binary relationship, of true or false value - to become the infocommunicational dispositif – with its immanent relations, potentially roll away to infinity, with overlapping and mutually interfering boundaries and with its belonging measurable in degrees. The use of mathematical and technological tools and languages, such as fuzzy logic and the semantic web, enabled the practical achievement of this approach, which was empirically applied in the domain of Brazilian popular songs. This complex domain, endowed with its own language (the song's language) has shown the inadequacy of traditional approaches based on metadata or collaborative classification methods and revealed to be successful by allowing a multi-dimensional approach of knowledge, considering it not as a finished product in itself, but emphasizing its dynamic character as process which the construction is of social and historical nature.

Keywords: Info communicational dispositif. Rhizome. Semiotics. Classification. OWL.

LISTA DE FIGURAS

Figura 1 – Tríplice fronteira entre Brasil, Argentina e Uruguai.....	16
Figura 2 – <i>Arbor naturalis et logicalis</i> , por Ramon Llull.....	24
Figura 3 – Sistemas de Organização do Conhecimento: estruturas e funções.....	25
Figura 4 – Arranjo APUPA (<i>Alien–Penumbral–Umbral–Penumbral–Alien</i>)	27
Figura 5 – Representação gráfica de um rizoma	35
Figura 6 – Esquematismo tensivo elaborado por Viggo Brøndal.....	54
Figura 7 – Quadrado das proposições, síntese das operações na lógica clássica	57
Figura 8 – Etapas de um processamento fuzzy, segundo o processo Mamdani.....	59
Figura 9 – Tipos de função de função de <i>fuzzification</i> : triangular; trapezoidal e Curva de Gauss.....	60
Figura 10 – Página web da “Yahoo!” em 1994.....	64
Figura 11 – Página do YouTube, vídeo para a canção “A Mulata é a Tal”	65
Figura 12 – Etapas de um sistema de organização do conhecimento dirigido pela representação (a) e pela recuperação (b).....	67
Figura 13 – Arquitetura da web semântica	70
Figura 14 – Resultados obtidos por tecnologias da web semântica em pesquisa no Google .	73
Figura 15 – Níveis de descrição na web semântica	74
Figura 16 – Tela do site Musicoverly	90
Figura 17 – Tela do site Aquela Música	91
Figura 18 – Grafo dos conjuntos de dados do projeto DBTune	93
Figura 19 – Tela de navegação do projeto <i>Linked Jazz</i>	96
Figura 20 – Tela do projeto Linked Jazz. Seleção de um artista.....	97
Figura 21 – Tipologia das canções segundo a Semiótica da Canção	101
Figura 22 – Análise da canção Querelas do Brasil, utilizando ferramenta Praat	107
Figura 23 – Análise da canção Aquarela do Brasil, utilizando ferramenta Praat.....	113
Figura 24 – Comparação das análises das canções Querelas do Brasil (a) e Aquarela do Brasil (b)	113
Figura 25 – Ilustração visual da descrição de canção e compositor, expressa em RDF.....	115
Figura 26 – Ilustração da descrição em RDF com instâncias, para a canção Querelas do Brasil e seus compositores	116
Figura 27 – Definição de propriedades fuzzy para expressar os graus de pertinência aos perfis melódicos	121

Figura 28 – Visão rizomática do sistema de organização do conhecimento construído	125
Figura 29 – Visão rizomática do sistema a partir da recuperação de dados relativos a Radamés Gnattali.....	126
Figura 30 – Interface web. Navegação pela classe Canção a partir de “Americanos”	127
Figura 31 – Interface web. Navegação pela relação “Compositor de”, explicitando a regra “Parceria”	128
Figura 32 – Interface web. Navegação pelas relações desdobradas a partir de “Americanos”	129
Figura 33 – Interface web. Navegação explicitando os graus de pertinência dos perfis melódicos das gravações	130

LISTA DE CÓDIGOS FONTE COMPUTACIONAIS

Código fonte 1 – Dados sobre a cidade Rio de Janeiro, expressa em XML	74
Código fonte 2 – Dados de cidade e fundador, expressos em RDF.....	75
Código fonte 3 – Instâncias da cidade Rio de Janeiro e dados da sua fundação, em RDF	76
Código fonte 4 – Regra que define um Fundador, expresa em OWL.....	78
Código fonte 5 – Definição de canção e compositor, expressa em RDF.....	114
Código fonte 6 – Definição das instâncias da canção Querelas do Brasil em RDF	116
Código fonte 7 – Descrição dos dados de canção, gravação, composição e intérprete, expressos em OWL.....	118
Código fonte 8 – Declaração da regra parceria, em OWL	119
Código fonte 9 – Definição de dados ligados, promovendo a integração com recursos do projeto DBPedia	120
Código fonte 10 – Definição de propriedades fuzzy para expressar os graus de pertinência aos perfis melódicos, em OWL	122

LISTA DE QUADROS

Quadro 1 – Exemplo de metadados para a canção “A Mulata é a Tal”	71
Quadro 2 – Metadados da <i>tag</i> ID3, versão 1.....	85
Quadro 3 – Metadados da <i>tag</i> ID3, versão 2.....	86
Quadro 4 – Metadados da canção Querelas do Brasil	104
Quadro 5 – Metadados da gravação de Querelas do Brasil, por Maurício Tapajós e Aldir Blanc	104
Quadro 6 – Classes identificadas a partir da abordagem empírica.....	123
Quadro 7 – Propriedades identificadas a partir da abordagem empírica	123
Quadro 8 – Regras identificadas a partir da abordagem empírica	124

LISTA DE ABREVIATURAS E SIGLAS

APUPA	Alien Penumbral Umbral Penumbra Alien
BIT	Binary Digit Dígito binário
CERN	Centre Européen pour la Recherche Nucléaire Centro Europeu para a Pesquisa Nuclear
FOAF	Friend-Of-A-Friend schema
ICSU	International Council for Science Conselho Internacional para a Ciência
ISMIR	International Society of Music Information Retrieval Sociedade Internacional de Recuperação de Informação em Música
ISKO	International Society for Knowledge Organization Sociedade Internacional de Organização do Conhecimento
KOS	Knowledge Organization System Sistema de Organização do Conhecimento
LDP	Linked Data Platform Plataforma de Dados Ligados
MPB	Música Popular Brasileira
MP3	MPEG-1/2 Audio Layer 3
OWL	Web Ontology Language Linguagem de Ontologia para Web
RDF	Resource Description Framework Framework para Descrição de Recursos
RISM	Répertoire International des Sources Musicales Repertório Internacional de Fontes Musicais
SESC	Serviço Social do Comércio
UNESCO	United Nations Educational, Scientific and Cultural Organization Organização das Nações Unidas para a Educação, a Ciência e a Cultura
UNISIST	United Nations International Scientific Information System Sistema de Informações Tecnológicas das Nações Unidas
URI	Uniform Resource Identifier Identificador Uniforme de Recurso
URL	Uniform Resource Locator Localizador de Recurso Uniforme
W3C	World Wide Web Consortium Consórcio World Wide Web
WWW	World Wide Web
XML	eXtensible Markup Language Linguagem de Marcação Estendida

SUMÁRIO

NOTA PRELIMINAR: ESCOLHAS TERMINOLÓGICAS E FONTES	
	BIBLIOGRÁFICAS..... 15
1	INTRODUÇÃO..... 16
1.1	JUSTIFICATIVA..... 17
1.2	PROBLEMA E QUESTÕES DE PESQUISA 18
1.3	OBJETIVOS 19
1.4	PREMISSAS E HIPÓTESES 19
1.5	ESTRUTURA DA TESE 21
2	CONTEXTUALIZAÇÃO TEÓRICA 23
2.1	ANTECEDENTES E CONSEQUENTES PARA UMA ABORDAGEM À COMPLEXIDADE 24
2.2	REPRESENTAÇÃO SIMBÓLICA DO CONHECIMENTO 30
2.3	TEORIA DO PENSAMENTO COMPLEXO 32
2.4	DISPOSITIVO COMO UNIDADE BÁSICA DO CONHECIMENTO..... 36
3	OBJETO DE PESQUISA: CANÇÃO POPULAR BRASILEIRA 40
3.1	CANÇÃO POPULAR ENQUANTO DOMÍNIO AUTÔNOMO 41
3.2	CANÇÃO POPULAR ENQUANTO DOMÍNIO COMPLEXO 42
3.3	CANÇÃO POPULAR ENQUANTO DISPOSITIVO INFOCOMUNICACIONAL 45
4	ASPECTOS METODOLÓGICOS RELACIONADOS À ORGANIZAÇÃO DO CONHECIMENTO..... 48
4.1	MÉTODO COMPLEXO 48
4.2	ESQUEMATISMO TENSIVO..... 50
4.3	LÓGICA FUZZY..... 56
4.4	COMPUTAÇÃO, WEB E DOCUMENTO 62
4.5	ORGANIZAÇÃO DO CONHECIMENTO DIRIGIDA PELA RECUPERAÇÃO. 67
4.6	WEB SEMÂNTICA..... 68
5	ASPECTOS METODOLÓGICOS RELACIONADOS À CANÇÃO POPULAR 82
5.1	GÊNERO MUSICAL..... 82
5.2	METADADOS PARA CANÇÕES POPULARES 84
5.3	MÉTODOS COLABORATIVOS..... 88
5.4	ABORDAGEM NA WEB 89
5.5	ABORDAGEM NA WEB SEMÂNTICA..... 93
5.6	ABORDAGEM SEMIÓTICA / TENSIVA..... 98
6	CAMPO EMPÍRICO – APLICAÇÃO E RESULTADOS 103
6.1	<i>CORPORA</i> DE ANÁLISE..... 103
6.2	ANÁLISE A PARTIR DE QUERELAS DO BRASIL..... 104
6.3	CONSTRUÇÃO DO SISTEMA E RESULTADOS ALCANÇADOS 114
7	CONCLUSÃO 131
	REFERÊNCIAS..... 138

REFERÊNCIAS BIBLIOGRÁFICAS	138
REFERÊNCIAS DISCOGRÁFICAS	147

NOTA PRELIMINAR: ESCOLHAS TERMINOLÓGICAS E FONTES BIBLIOGRÁFICAS

Esta tese apresenta escolhas terminológicas que devem ser esclarecidas preliminarmente, no intuito de facilitar sua compreensão.

O primeiro esclarecimento diz respeito à adoção da expressão “abordagem a canções populares”. A noção de “abordar” é central para a teoria do pensamento complexo aqui adotada e que será discutida no curso da tese. Entretanto, desde o título é importante destacar que esta teoria considera que domínios complexos, como é o caso das canções populares, não são “solucionáveis”, mas sim “abordáveis”, ou seja, que não há a pretensão de fornecer uma resposta irrefutável nem tampouco de construir um sistema de organização do conhecimento fechado e definitivo, mas sim identificar e viabilizar diferentes perspectivas de se abordar o problema.

O segundo esclarecimento está relacionado ao uso de uma palavra em língua inglesa: *fuzzy*. Em tradução literal, *fuzzy* quer dizer “frisado, esfiapado, penugento”, e, por extensão, e também no contexto científico, é usualmente traduzido como “difuso, nebuloso, impreciso, vago” (MICHAELIS, 2002, p. 133). O emprego destas traduções na literatura científica em língua portuguesa não é consensual nem consistente (WEBER; KLEIN, 2003). Considerando-se que haveria prejuízo semântico ao optar arbitrariamente por alguma das traduções, decidiu-se, adotar o termo original em inglês e grafá-lo sem o destaque em itálico.

O terceiro esclarecimento diz respeito à utilização do termo “conceito” exclusivamente para se referir à “unidade básica do conhecimento”, conforme definido por Dahlberg (1978), (2014). Para a aceção de uso comum, serão adotados os termos noção ou ideia, conforme o contexto.

O último esclarecimento terminológico diz respeito a uma deliberada omissão: a decisão de não adotar a sigla SOC para “Sistema de Organização do Conhecimento”, devido à sua cacofonia (soa como a conjunção adversativa “só que”).

Um esclarecimento geral diz respeito às fontes bibliográficas e suas traduções. Sempre que possível, buscou-se utilizar os textos originais dos autores, realizando traduções livres e fazendo constar em notas de rodapé o conteúdo original. Nos casos em que não foi possível obter o texto na língua original, recorreu-se a versões já traduzidas para o português. No caso específico de obras originais de Viggo Brøndal e de Ernst Cassirer escritas, respectivamente, em dinamarquês e em alemão, recorreu-se a traduções em língua francesa devido ao desconhecimento deste autor em relação àqueles idiomas.

1 INTRODUÇÃO

As fronteiras são componentes fundamentais de qualquer sistema de organização do conhecimento. Elas são habitualmente utilizadas com a função de delimitação, promovendo a separação entre duas ou mais entidades de forma inequívoca e integral.

Um exemplo corrente é sua aplicação na noção de fronteira geográfica. A Figura 1 apresenta uma imagem de satélite da região da tríplice fronteira entre o Brasil, a Argentina e o Uruguai.

Figura 1 – Tríplice fronteira entre Brasil, Argentina e Uruguai.



Fonte: Adaptado de Google Mapas ¹, com inserção dos nomes dos países.

A característica de inequívocidade da fronteira permite determinar se uma pessoa é brasileira, argentina ou uruguaia de acordo com o lado da fronteira em que ela nasceu. A característica da integralidade da fronteira não permite outra opção nem uma combinação entre as opções. No exemplo citado, se uma pessoa é uruguaia, ela é integralmente uruguaia, nunca “um pouco uruguaia” nem “mais ou menos brasileira”.

A caracterização das fronteiras como rígidas e mutuamente exclusivas constitui o princípio básico de funcionamento dos sistemas de organização do conhecimento tradicionais:

¹ Google Mapas. Disponível em: <<http://maps.google.com>>. Acesso em: ago 2015.

a dicotomia, ou seja, a separação sem sobreposições nem ambiguidades (JACOB, 2004, p. 520–522).

Contudo, estudos das ciências sociais e humanas a partir do século XX vem considerando que as fronteiras também podem ser entendidas como lugares de integração, de confluência, regidos por princípios de relatividade e de multiplicidade, constituindo antes um “traço de união” que apenas de separação (DELEUZE; GUATTARI, 1980). No exemplo da região da tríplice fronteira, existe um designativo para quem pertence a esta região de integração: o gaúcho.

Esta tese propõe uma concepção de Organização do Conhecimento que considere as fronteiras nesta segunda acepção – como lugar de integração e de sobreposição. Para tanto, é discutida a pertinência desta abordagem a partir de uma revisão histórica que busca analisar seus antecedentes à luz dos estudos contemporâneos na área. Em seguida são investigadas as bases teóricas e metodológicas adequadas à formulação de um sistema de organização do conhecimento, cujos resultados são discutidos a partir de sua aplicação no domínio das canções populares.

1.1 JUSTIFICATIVA

Os sistemas de organização do conhecimento tradicionais, ou seja, aqueles que operam com base na dicotomia, tais como as tabelas de classificação, as taxonomias e os tesouros, são necessários em muitas aplicações. Retomando o exemplo da fronteira geográfica e considerando as suas implicações em um sistema geopolítico, seria inconcebível para um governo que as fronteiras não fossem rígidas e mutuamente exclusivas. De outra forma, um gaúcho votaria na eleição presidencial de qual país? Ou seria regido por qual Constituição?

Entretanto, há situações nas quais as fronteiras não são claras nem inequívocas. Dois casos exemplares são o problema do ornitorrinco nas taxonomias biológicas e a recorrente existência de uma entrada “Outros” nas tabelas de classificação. Estes dois exemplos revelam que, para lidar com o que não é adequado às fronteiras dicotômicas, os sistemas de organização do conhecimento tradicionais precisam adotar estratégias de simplificação, como colocar o que é ambíguo na entrada “Outros” ou considerar o ornitorrinco como mamífero ²,

² Estudos genômicos afirmam que o ornitorrinco é simultaneamente mamífero, ave e réptil. Ver: WARREN, W. et al. Genome analysis of the platypus reveals unique signatures of evolution. In: *Nature*, v.453, n.7192, p.175-183, May, 8, 2008.

adaptando o “problema” a um sistema já existente, neste último caso, a taxonomia dos seres vivos. Em muitos domínios do conhecimento, especialmente naqueles de âmbito social e cultural, as fronteiras são naturalmente sobrepostas e o que seria considerado como exceção em uma visão tradicional (o ornitorrinco) passa a ser tão ou mais frequente que as regras (MAZZOCCHI, 2006).

Retomando novamente o exemplo da fronteira geográfica, mas agora considerando a necessidade de uma organização de viés cultural, tal como a categorização³ dos trajes ou da gastronomia, é mais razoável considerar como pertencentes a uma mesma categoria o gaúcho brasileiro e os *gauchos* uruguaios e argentinos que o gaúcho brasileiro e o potiguar, por exemplo.

A noção de pertencimento está intimamente ligada à ideia de gradação, da possibilidade de expressar o quanto algo pertence ou não a uma dada categoria. Nos sistemas tradicionais, tendo em vista o princípio da integralidade, o pertencimento é reduzido a uma relação binária (valores “verdadeiro” ou “falso”), ou seja, de pertencimento integral ou de não pertencimento em absoluto.

A necessidade de organizar o conhecimento de maneira não dicotômica e para além do paradigma simplificador justifica a proposta de elaboração de uma concepção diferente de sistema de organização do conhecimento, que seja adequado para lidar com domínios que são caracterizados pelas simultaneidades e pelas gradações: os domínios complexos.

1.2 PROBLEMA E QUESTÕES DE PESQUISA

Esta tese trata do problema da organização do conhecimento em domínios complexos, aqueles caracterizados pela sobreposição das fronteiras e pelas relações de pertencimento mensuráveis em graus, que não são alcançáveis nos sistemas de organização do conhecimento tradicionais.

As questões de pesquisa levantadas a partir deste problema podem ser divididas em duas partes complementares:

³ A distinção entre os pares de termos “classe/classificação” e “categoria/categorização” adotada neste ponto segue o princípio estabelecido por Elin K. Jacob em “*Classification and Categorization: a difference that makes a difference*” (2004). Sucintamente, segundo a autora “classificação” é um processo de atribuição de valores mutuamente exclusivos (as classes), dentro de um sistema hierárquico previamente ordenado; enquanto a categorização é um processo de base cognitiva que consiste em dividir as entidades em grupos (as categorias) pelo compartilhamento de aspectos de similaridade entre seus membros. Como esta distinção não é consensual e será sobreposta por outras formulações teóricas no curso desta tese, a discussão a respeito não será prolongada.

- A primeira, de fundo epistemológico, que solicita uma abordagem adequada para lidar com o complexo em substituição ao modelo simplificador tradicional dos sistemas de organização do conhecimento;
- A segunda, de fundo aplicado, relacionada aos requisitos técnicos e operacionais para construção e utilização de um sistema de organização do conhecimento adequado ao paradigma complexo.

1.3 OBJETIVOS

O objetivo geral desta pesquisa é apresentar as condições adequadas para a elaboração de um sistema de organização do conhecimento capaz de abordar domínios complexos.

Para a consecução deste objetivo geral, são identificados os seguintes objetivos específicos:

- Estabelecer as bases teóricas para abordagem a domínios complexos.
- Elicitar os requisitos para abordar a complexidade em sistemas de organização do conhecimento.
- Determinar as técnicas e ferramentas adequadas para construção de sistemas de organização do conhecimento aplicável a domínios complexos.
- Construir empiricamente um sistema de organização do conhecimento, através da aplicação das técnicas e ferramentas adequadas a um domínio complexo - as canções populares brasileiras.

1.4 PREMISSAS E HIPÓTESES

Esta tese parte de três premissas que, embora não possam ser consideradas axiomáticas, embasam a proposta apresentada e permitem que seja mantido o fio condutor das hipóteses apresentadas a seguir.

- A primeira premissa é a da arbitrariedade dos sistemas de organização do conhecimento. Derek Wilton Langridge (1925–2000) afirma que “toda classificação está relacionada a um propósito” e que “os mesmos objetos podem ser classificados de maneiras diferentes dependendo do propósito a que se destinam” (LANGRIDGE, 1977, p.21), como foi exemplificado anteriormente ao aludir à diferença entre se considerar os habitantes da região

da tríplice fronteira pelo propósito de determinar sua nacionalidade ou pelo propósito de agrupá-los de acordo com critérios culturais. É importante, no entanto, não confundir essa premissa com a característica da multidimensionalidade, desejável aos sistemas de organização do conhecimento. Considerar que o conhecimento deve poder ser abordado por múltiplas dimensões (multidimensionalidade), não significa desconsiderar que estas abordagens (individualmente, ou mesmo quando combinadas) são arbitrárias, fruto de escolhas que são determinadas de acordo com o propósito do sistema elaborado.

- A segunda premissa é decorrente da anterior e diz respeito à natureza dos sistemas de organização do conhecimento. Estes sistemas são instrumentos de poder, arbitrários na sua gênese (como afirma Langridge) e enviesados na sua aplicação. Segundo Pierre Bourdieu (1930-2002), os sistemas de classificação constituem a doxa de um campo e são simultaneamente estruturados (pelas relações de poder que os determinam) e estruturantes (das mesmas relações de poder e do próprio campo) (BOURDIEU, 1989), do que resulta que não pode haver ingenuidade de considerar sua construção ou utilização sob uma perspectiva neutra ou imparcial.
- A terceira premissa desta tese considera que a gênese e a determinação semântica das noções adotadas estão postas, e que seu entendimento é compartilhado dentro da abrangência do sistema de organização do conhecimento em questão. Novamente recorrendo ao exemplo inicial, esta premissa diz que “argentino”, “brasileiro” e “uruguaio” são noções que existem e que são compreensíveis para os utilizadores de um sistema de organização do conhecimento por “nacionalidade” (que também é uma noção assentada). Esta premissa tem uma consequência importante que é a de afastar esta tese de qualquer inclinação a discussões de uma Teoria do Conhecimento, sobretudo aquelas de caráter cognitivista.

Consideradas as três premissas, é possível enunciar as hipóteses desta pesquisa.

A hipótese analítica desta tese é que é possível elaborar sistemas de organização do conhecimento capazes de abordar domínios complexos, considerando-se a multidimensionalidade do conhecimento a partir de fronteiras sobrepostas e relações de pertencimento expressas em graus.

Esta hipótese analítica compatibiliza as hipóteses de trabalho, enunciadas a seguir:

- Artefatos teóricos específicos, como os dispositivos infocomunicacionais, permitem abordar e representar a complexidade do conhecimento;
- Princípios operativos, como o esquematismo tensivo, permitem a manipulação e sistematização de domínios complexos do conhecimento;
- Novos recursos computacionais, como a web semântica e os sistemas fuzzy, permitem construir sistemas de organização do conhecimento capazes de lidar com o paradigma complexo.

1.5 ESTRUTURA DA TESE

A presente tese está organizada da seguinte forma:

No capítulo 2 é apresentada a contextualização teórica que fundamenta as discussões desenvolvidas ao longo do trabalho. Esta fundamentação é realizada buscando traçar os antecedentes históricos do pensamento complexo e suas consequências para a Organização do Conhecimento contemporânea. Em seguida são discutidos os princípios da representação simbólica do conhecimento e da teoria do pensamento complexo. Encerrando a contextualização teórica, são retomadas as noções de dispositivo infocomunicacional e de documento, enfatizando a sua centralidade para a presente pesquisa.

No capítulo 3 é discutida a especificidade da canção popular brasileira. A partir de um metaexemplo (uma canção popular específica), são exploradas as características que definem a canção brasileira como um domínio autônomo, dotado de uma linguagem própria, e, simultaneamente, como um domínio complexo. Por fim, é realizada a aproximação entre as noções de dispositivo infocomunicacional e canção popular, configurando esta como um tipo exemplar de dispositivo.

No capítulo 4 são discutidos os aspectos metodológicos para o sistema de organização do conhecimento proposto, incluindo o método para abordagem da complexidade, o esquematismo tensivo e a lógica fuzzy. Além disso, são discutidas as noções de web, da organização do conhecimento dirigida pela recuperação e são apresentados os princípios de funcionamento da web semântica.

No capítulo 5 são discutidos os aspectos metodológicos relacionados ao domínio musical e, especialmente, as suas implicações para o domínio das canções populares. São abordadas as metodologias de classificação em gêneros musicais, a descrição por metadados, as abordagens colaborativas e as metodologias relacionadas aos ambientes da web e da web semântica, bem como as abordagens semióticas.

No capítulo 6 é apresentada a aplicação do referencial teórico e metodológico para o domínio das canções populares brasileiras e são discutidos os resultados desta abordagem.

No capítulo 7 é apresentada a conclusão e são discutidos alguns possíveis desdobramentos desta pesquisa através do apontamento de trabalhos futuros.

2 CONTEXTUALIZAÇÃO TEÓRICA

Esta seção apresenta de forma linear um percurso que se fez a partir de cruzamentos e circularidades. A intenção deste capítulo não é elaborar um inventário completo das possibilidades teóricas, mas sim apresentar os autores selecionados e contextualizar as escolhas realizadas, apresentando-as de maneira coerente aos objetivos anteriormente enunciados.

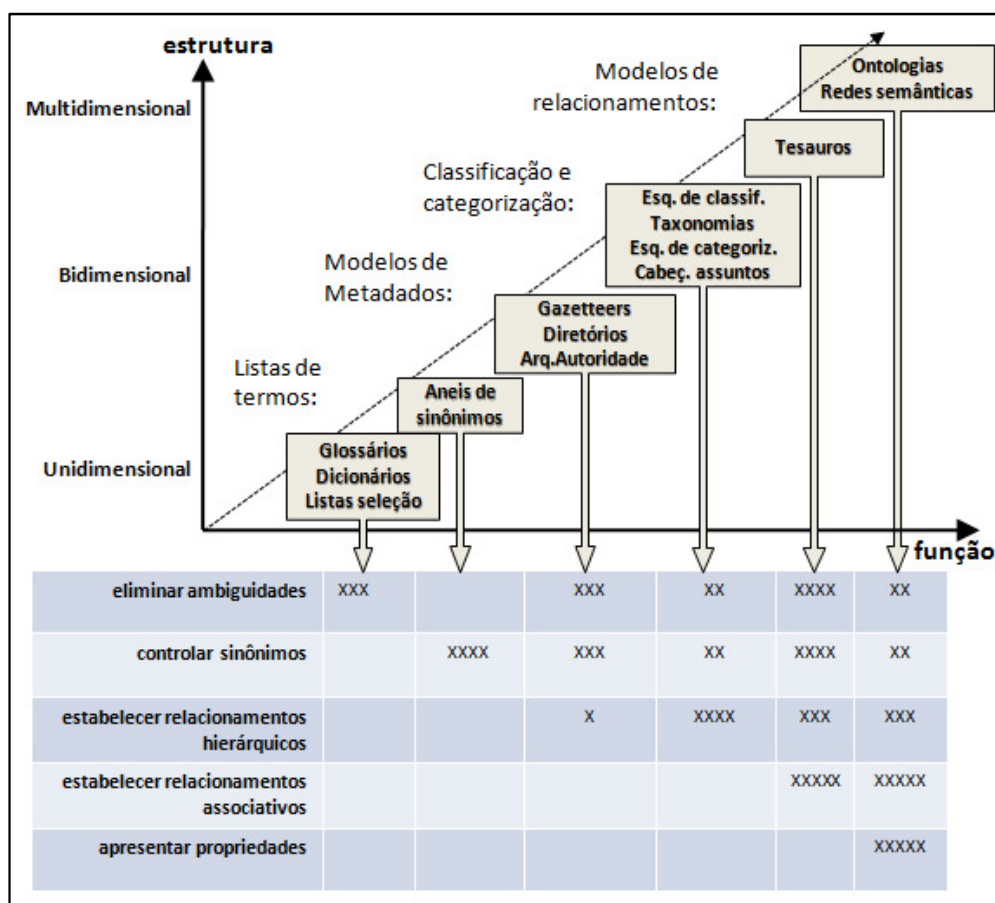
Esta contextualização teórica é dividida em quatro partes, destacando-se os aportes dos seguintes autores:

- Uma breve revisão histórica que recupera alguns antecedentes do pensamento complexo, especialmente nas obras de Shiyali Ramamrita Ranganathan e de Georges-Louis Leclerc (conde de Buffon), confrontando-os com a tipologia atual dos sistemas de organização do conhecimento e discutindo-os com as propostas de novas abordagens destes sistemas, defendidas por pesquisadores contemporâneos como Claudio Gnoli e Rick Szostak.
- Uma visão do conhecimento enquanto representação simbólica, dentro de uma perspectiva semiótica, a partir da filosofia das formas simbólicas de Ernst Cassirer. Esta perspectiva é contextualizada a partir do estudo sobre representação do conhecimento, de Maria Nélida González de Gómez, e discutida no âmbito da Organização do Conhecimento a partir de trabalhos de Brian Vickery e de Rosali Fernandez de Souza.
- Uma apresentação da Teoria do Pensamento Complexo formulada por Edgar Morin, considerando os aportes relativos aos sistemas complexos a partir de Roberto Poli e a noção de fronteira em Gilles Deleuze.
- A retomada da noção de “documento” e a aproximação da noção de “dispositivo infocomunicacional” na Organização do Conhecimento, a partir dos trabalhos de Regina Marteleto e Viviane Couzinet e dos precursores destas noções (respectivamente, Paul Otlet para “documento” e Michel Foucault para “dispositivo”).

- a exclusividade mútua: um elemento não pode pertencer a mais de uma classe simultaneamente;
- a teleologia: cada classe atende objetivamente a um propósito específico, ou seja, representa um determinado aspecto objetivo da realidade;
- a hierarquia: as classes são dispostas seguindo a uma estrutura de subordinação e de ordenação rígidas, determinadas no momento da construção do sistema de classificação.

Os principais tipos de sistemas de organização do conhecimento utilizados contemporaneamente são baseados em estruturas hierárquicas e dicotômicas e, usualmente, a sua eficácia é avaliada precisamente em função do quanto estes sistemas conseguem estruturar hierarquicamente e separar dicotomicamente os elementos na busca de representar objetivamente o mundo real, conforme ilustra a Figura 3 (ZENG, 2008).

Figura 3 – Sistemas de Organização do Conhecimento: estruturas e funções



Fonte: Adaptado e traduzido de (ZENG, 2008, p. 161).

Dos treze tipos atuais de sistemas de organização do conhecimento elencados por Zeng, apenas os “anéis de sinônimos”, por razões óbvias, não tem por função eliminar ambiguidades (e sim, evidenciá-las). E apenas os sistemas dos primeiros grupos (como as “listas de termos”, os “glossários/dicionários” e os próprios “anéis de sinônimos”) não são hierarquicamente estruturados. A evolução dos sistemas de organização do conhecimento e sua metaclassificação em relação às funcionalidades oferecidas e à possibilidade de representação multidimensional do conhecimento, está, portanto, diretamente relacionada ao paradigma clássico.

No entanto, a necessidade de conceber diferentes possibilidades para organização do conhecimento, para além das estruturas hierárquicas e dicotômicas, não é absolutamente uma novidade.

À proposta de uma taxonomia biológica dos seres vivos defendida por Carl von Linné (1707-1778) em *Systema Naturae* (1736) e em *Genera Plantarum* (1737), Georges-Louis Leclerc (1707-1788), nobilitado conde de Buffon, contrapôs em *Histoire Naturelle* (1749), o argumento de ser impossível refletir a realidade da natureza, pois “esta caminha por gradações desconhecidas e, por consequência, ela não se pode prestar totalmente a essas divisões [referia-se às classes de gênero/espécie defendidas por von Linné]” (POMBO, 1998, p. 9)^{4 5}.

Portanto, há mais de três séculos, e quando provavelmente ainda nem se conhecia o ornitorrinco⁶, Buffon já apontava a inadequação do princípio da integralidade, denunciada pela existência de uma gradação, cuja operação, entretanto, era para ele desconhecida.

Tratando da organização de bibliotecas, mais especificamente da disposição física dos livros em estantes segundo uma classificação por assuntos, Shiyali Ramamrita Ranganathan (1892-1972) enunciou a necessidade de um modelo que permitisse representar as “relações de vizinhança”⁷ entre os assuntos e propôs o arranjo APUPA (acrônimo para *Alien–Penumbral–Umbral–Penumbral–Alien*), uma técnica para tratar as fronteiras entre os assuntos de uma biblioteca considerando as suas relações de proximidade (RANGANATHAN, 1967, pt. Q, chap. QB, p. 383-385).

⁴ Grifos nossos.

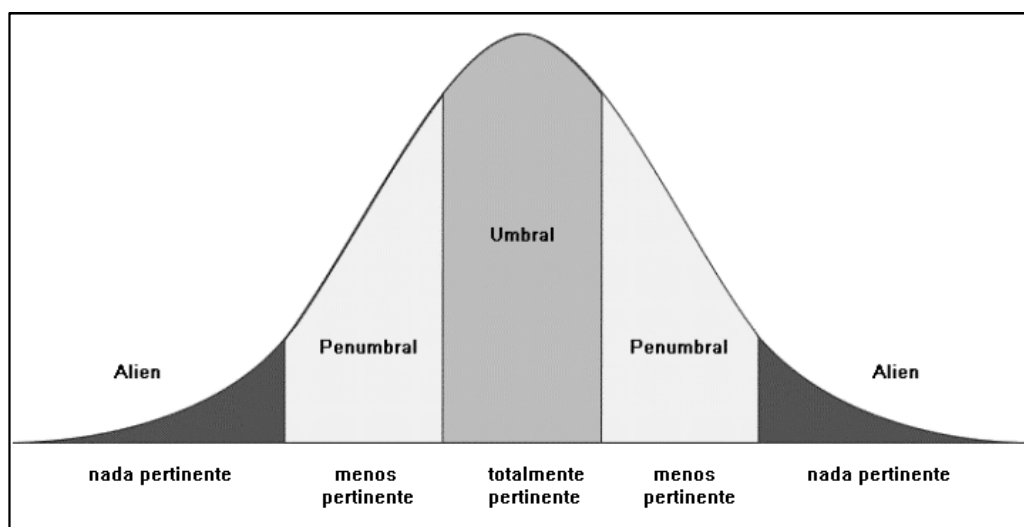
⁵ Para uma leitura pormenorizada da disputa teórica entre von Linné e Buffon, ver “*Da classificação dos seres à classificação dos Saberes*”, de Olga Pombo (1989). Para as consequências desta disputa e sua influência sobre o evolucionismo de Charles Darwin e posteriormente sobre as teorias eugênicas, ver VIGNAUX, Georges. *O Demônio da classificação: pensar / organizar*. Lisboa: Instituto Piaget, 2000.

⁶ O primeiro registro conhecido de descrição do ornitorrinco data de 1799, e foi feito pelo zoólogo George Shaw. SHAW, G.; NODDER, F.P. *Platypus anatinus*. The Duck-billed Platypus. In: *The Naturalists' Miscellany*, London: F.P. Nodder & Co, v.10, n.8, p. 385-386, 1799.

⁷ No original, em inglês, “*neighborhood-relations*”.

O princípio de operacionalização da APUPA propõe colocar no centro (*Umbral*) os livros mais pertinentes a um determinado assunto. Os livros de importância marginal, mas ainda pertinentes ao assunto em questão, vão sendo colocados ao lado (*Penumbral*); e, de forma decrescente, se afastando do centro, colocam-se os livros menos pertinentes até chegar àqueles totalmente desconectados do assunto (*Alien*). O arranjo APUPA é ilustrado pela Figura 4.

Figura 4 – Arranjo APUPA (*Alien–Penumbral–Umbral–Penumbral–Alien*)



Fonte: Adaptado e traduzido de (VELMURUGAN, 2010).

Ao explicar a proposta da APUPA, Ranganathan afirma não ser possível executá-la na prática, pois o arranjo físico só permite uma configuração por vez, e seria impossível refletir as diversas possibilidades de relações entre os múltiplos assuntos existentes em uma biblioteca. Ao enunciar as limitações da APUPA, Ranganathan afirma que se trata de um problema “de complexidade” para o qual, naquele momento, não havia ferramentas capazes de resolver (ibidem).

Ranganathan enuncia, portanto, um problema de complexidade, e enfatiza a questão da delimitação e da sobreposição das fronteiras e, com clareza, afirma que nos domínios complexos, a organização deveria se originar do centro (do *Umbral*) e não dos limites (a partir das fronteiras). Sem fazer menção a Ranganathan, Edgar Morin, ao propor um método para lidar com a complexidade afirma precisamente a mesma coisa: “Os conceitos não se definem jamais por suas fronteiras, mas a partir de seu núcleo [...]. As fronteiras são sempre fluidas, são sempre interferentes. Deve-se, pois, buscar definir o centro” (MORIN, 2005b, p. 72–73).

Pesquisas contemporâneas em Organização do Conhecimento vem afirmando a necessidade de se considerar a complexidade (e não mais evitá-la) e apontam os avanços da lógica matemática e da computação como ferramentas que possibilitariam a concepção de novos sistemas de organização do conhecimento:

Este surgimento de inovações teóricas a partir do século XX, diretamente ou não, tiveram também um impacto sobre o pensamento classificatório. Por exemplo, diferentes tipos de lógicas polivalentes foram formulados a partir da obra de pensadores inovadores, tais como Łukasiewicz, Gödel, e Zadeh.[...] Mais recentemente, a criação de hipertextos e da web trouxeram a difusão de uma cultura em rede, da organização da informação (e, talvez, de caminhos mentais).⁸ (MAZZOCCHI, 2013, p. 370. Tradução nossa)

A metáfora imagética do conhecimento sofreu uma significativa mudança. Em substituição à Árvore de Porfírio, são propostas imagens do conhecimento como mapas ou labirintos (ECO, 2013) e (FEDELI, 2013) ou como redes ou rizomas (MAZZOCCHI, 2013), (DELEUZE; GUATTARI, 1995).

Diante desta mudança paradigmática, pesquisadores da área de Organização do Conhecimento vem reconhecendo a insuficiência das visões disciplinares tradicionais, restritas a domínios específicos do conhecimento (DAHLBERG, 2014) e asseverando a necessidade de investigar bases teóricas mais adequadas para lidar com a multidimensionalidade do conhecimento (GOMES, 2009).

Na edição de 2007 da seção espanhola daquela que é provavelmente a mais representativa associação da área de Organização do Conhecimento, a ISKO (*International Society for Knowledge Organization*), os pesquisadores se reuniram para discutir os rumos da área e elaboraram propostas para o futuro imediato dos estudos em Organização do Conhecimento. Estas propostas foram discutidas coletivamente durante as seções de trabalho e registradas em um documento textual posteriormente intitulado “Manifesto de León”, e teve redação final Claudio Gnoli e Rick Szostak, sendo publicado na revista *Knowledge Organization* daquele mesmo ano. Dentre as propostas apresentadas, destaca-se:

- que a tendência atual de uma crescente interdisciplinaridade do conhecimento exige, essencialmente, novos sistemas de organização do conhecimento (KOS), com base em uma revisão dos fundamentos dos KOS tradicionais centrados em disciplinas;
- que esta inovação não é apenas desejável, mas também possível, e deve ser implementada desenvolvendo de fato alguns novos KOS;

⁸ No original, em inglês: “This grew out of groundbreaking theoretical findings from the XX century, directly or not, have had also an impact on classificatory thinking. For instance, different kinds of (polyvalent) logic have been formulated through the work of innovative thinkers such as Łukasiewicz, Gödel, and Zadeh. [...] More recently, the creation of hypertext and the web has brought about the diffusion of a net-styled culture, information organization (and perhaps mental pathways)”.

- que em vez de disciplinas, as unidades básicas dos novos KOS devem ser fenômenos do mundo real, tais como eles são representados no conhecimento humano.⁹ (GNOLI; SZOSTAK, 2007, p. 6. Tradução nossa. Grifos nossos.)

Esta tese se alinha duplamente à proposta de uma Organização do Conhecimento não disciplinar e que considera os avanços tecnológicos e as transformações trazidas por tais avanços para as atividades de criação e utilização de sistemas de organização do conhecimento.

Ao mencionar a aplicação em domínios complexos, a aceção de “domínio do conhecimento” adotada nesta tese remete antes à noção matemática – oriunda da Teoria dos Conjuntos – do que àquela postulada pela Análise de Domínio – formulada por Birger Hjørland e Hanne Albrechtsen (1995). Esta distinção se faz necessária uma vez que, na visão matemática, um domínio é o conjunto a partir do qual se operam as funções, que, por sua vez, definem o contradomínio e a imagem do domínio original. Em outras palavras, o ponto fulcral para a teoria matemática dos conjuntos consiste nas relações estabelecidas e não no domínio enquanto conjunto, isoladamente¹⁰. Já na teoria da Análise de Domínio, um domínio corresponde ao conjunto de descrições locais de uma parcela do conhecimento – restrita a uma utilização previamente contextualizada. Esta visão considera dois pressupostos: o primeiro afirma que os utilizadores de um sistema de organização do conhecimento baseado nesta teoria devem conhecer a maneira através da qual são informados; o segundo afirma que apenas descrições locais a este conjunto de utilizadores são possíveis (HJØRLAND, 2002). Neste sentido, a Análise de Domínios é contrária à ideia de sistemas de classificação de carácter universal ou mesmo de ferramentas de propósito geral, como é o caso da ferramenta de busca Google (que será abordada novamente no decorrer deste trabalho)¹¹.

⁹ No original, em inglês: “[...] - the current trend towards an increasing interdisciplinarity of knowledge calls for essentially new knowledge organization systems (KOS), based on a substantive revision of the principles underlying the traditional discipline-based KOS;

- this innovation is not only desirable, but also feasible, and should be implemented by actually developing some new KOS;

- instead of disciplines, the basic units of the new KOS should be phenomena of the real world as it is represented in human knowledge”.

¹⁰ Para detalhes sobre a Teoria dos Conjuntos na Matemática, ver CASTRUCCI, Benedito. Elementos de teoria dos conjuntos. Grupo de Estudos do Ensino da Matemática, GEEM, 1969.

¹¹ Foge ao âmbito deste trabalho problematizar a Análise de Domínio em relação a abordagens não disciplinares da Organização do Conhecimento. Para uma visão mais abrangente desta discussão, ver, dentre outros, a série de artigos críticos publicados por Rick Szostak e Birger Hjørland, na qual as abordagens são discutidas:

- HJØRLAND, Birger. Core classification theory: A reply to Szostak. *Journal of Documentation*, v. 64, n. 3, p.333-342, 2008.

- SZOSTAK, Rick. Interdisciplinarity and classification: a response to Hjørland. *Journal of documentation*, v. 64, n. 4, p. 479-481, 2008.

A proposta de uma Organização do Conhecimento não disciplinar, por sua vez, também não é inédita e remete ao legado do *Classification Research Group* britânico, conforme registrado nos boletins publicados regularmente entre 1952 e 1968 ¹² e confirmado em trabalhos individuais de alguns de seus membros, como Douglas Foskett (1918-2004) que questiona em “*Beyond reductionism*” (FOSKETT, 1970) a crença filosófica corrente que todas as atividades humanas poderiam ser representadas e explicadas através de respostas elementares simplificadas (o reducionismo científico); Derek Austin (1921-2001) que na obra sugestivamente intitulada “*Two steps forward*” (1971), lança as bases para a concepção de sistemas de organização do conhecimento em ambientes mecanizados; e, mais recentemente, reafirmadas nos últimos trabalhos de Brian Campbell Vickery (1918-2009), como em “*Information Representation and Retrieval in the Digital Age*” (2004), e “*Faceted classification for the web*” (2008) nos quais o autor aborda com lucidez as necessidades de transformação das concepções de representação e recuperação frente às novas tecnologias. Além da vasta contribuição à epistemologia da área, Vickery foi também um importante responsável pela sistematização da base simbólica dos sistemas de organização do conhecimento, que é discutida a seguir.

2.2 REPRESENTAÇÃO SIMBÓLICA DO CONHECIMENTO

Considerando o conhecimento como “a relação do pensamento com o real”, Maria Nélida González de Gómez identificou três grandes momentos do conhecimento no quadro da cultura ocidental: o momento ontológico – quando a excelência do conhecer não passa pelo representar –; o momento gnosiológico – quando conhecer é representar e o mundo só é quando representado –; e o momento semiótico – quando o conhecimento é autônomo, independente do sujeito e do objeto (GONZÁLEZ DE GÓMEZ, 1993, p. 218).

Quando a unidade entre o homem e o mundo que caracterizava o momento ontológico foi gradativamente sendo substituída pela unidade da consciência, o conhecimento passou a ser representação, constituindo o momento gnosiológico, no qual “conhecer é representar e o

- SZOSTAK, Rick. Comment on Hjørland's concept theory. *Journal of the American Society for Information Science and Technology*, v. 61, n. 5, p. 1076-1077, 2010.

Hjørland, Birger. Answer to Professor Szostak (concept theory). *Journal of the American Society for Information Science and Technology*, v. 61, n. 5, p. 1078-1080, 2010.

¹² Neste período foram publicados os nove primeiros números do *Classification Research Group Bulletin*. Com significativas modificações em relação aos integrantes do grupo original e com maior intervalo de tempo foram posteriormente publicados mais três números: o número 10 em 1973, o número 11 em 1978, e o número 12 já em 1985.

mundo deixa de ser a totalidade do contexto da experiência para ser o lugar da representação” (ibidem, p. 219).

René Descartes (1596-1650) fundou no século XVII o que veio a se converter no paradigma científico predominante da cultura ocidental desde então: a separação entre o sujeito pensante (*ego cogitans*) e a coisa entendida (*res extensa*). Este paradigma estabelece um princípio de verdade baseado na descoberta de uma “ordem perfeita” na realidade. González de Gómez nos indica que foi no século XIX que teve início um movimento em sentido inverso ao do momento gnosiológico que se intensificaria no século XX: “o sujeito perderá toda força explicativa – não só na esfera do conhecimento, mas também enquanto agente de ações e transformações sociais.” (ibidem, p.221). A autora destaca ainda que o estruturalismo, os novos estudos da semiótica e as epistemologias sem sujeito, passaram a considerar a linguagem como a base para o conhecimento.

O filósofo Ernst Cassirer (1874-1945) desenvolveu uma filosofia da cultura humana, tomando por base a concepção relacional do conhecimento a partir da linguagem:

A linguagem, nas suas designações de conteúdos e relações espaciais, possui tal esquema, com o qual ela precisa, necessariamente, relacionar todas as representações intelectuais para, assim, torná-las apreensíveis e representáveis pelos sentidos. É como se todas as relações intelectuais e ideais somente pudessem ser apreendidas pela consciência linguística no momento em que a linguagem as projeta no espaço e nele as ‘copia’ analogicamente. (CASSIRER, 2001a, p. 211–212)

Para Cassirer, portanto, relações das representações se dão entre as percepções e a inteligência, com a segunda organizando o mundo graças às primeiras. A consequência deste deslocamento foi novamente inserir o sujeito na estrutura, conciliando, a partir de uma fenomenologia do conhecimento, a estrutura objetiva da realidade com o sujeito cognoscente. A esta conciliação correspondem três camadas de formas simbólicas através das quais o homem pensa o mundo:

(I) a de atividade quando sentimos, agimos e, desta forma, aprendemos as coisas;
(II) a da representação intuitiva quando a partir destas apreensões, temos percepções do mundo;
(III) e por fim, a da *significação* quando, graças a estas percepções, elaboramos para nós próprios os sentidos ou conceitos do mundo e aplicamo-los ao mesmo. São as representações. (VIGNAUX, 2000, p. 30–31).

Em suma, para Cassirer, são as representações do mundo que constroem a própria realidade (CASSIRER, 1985). Segundo Pierre Bourdieu (declaradamente influenciado pela leitura de Cassirer na construção de sua teoria do poder simbólico), foi Cassirer quem originalmente concebeu o entendimento das representações simbólicas como “instrumentos de

conhecimento” e de construção do mundo objetivo a partir de estruturas subjetivas (BOURDIEU, 1989, p. 8–16).

Esta concepção é fundamental para a proposta de um tipo de sistema de organização do conhecimento capaz de abordar a complexidade, por permitir superar o princípio teleológico inerente à abordagem clássica, realizando-a a partir de uma representação simbólica do conhecimento, na qual também as fronteiras não são mais consideradas apriorísticas, mas como processos de construção social e histórica, produzidas simbolicamente (DELEUZE, 1968).

Brian Vickery afirma que “a representação do conhecimento na forma simbólica é uma questão que preocupava o mundo da documentação desde sua origem”. E que, com o avanço das tecnologias, além dos documentos e índices físicos, as representações simbólicas passaram a ser uma preocupação para as bases de dados, para os programas de computador; para a inteligência artificial e para a construção de modelos cognitivos, e alerta para o fato que “em todos estes campos é necessário decidir como o conhecimento pode ser representado de forma que as representações possam ser manipuladas” ¹³ (VICKERY, 1986, p. 145. Tradução nossa).

A preocupação de Vickery parece ser a mesma de Rosali Fernandez de Souza quando a pesquisadora afirma que a representação para fins de recuperação é “de grande valia na busca de melhores soluções operacionais e de ordem prática”, mas alerta para que não se descuide das “origens do processo de classificação” e da base teórica, asseverando que “frente a complexidade da questão, deva haver uma reflexão mais consciente” (SOUZA, 2000, p. 4).

Um ponto de partida para atender a esta necessidade de reflexão mais consciente pode estar na própria origem latina do termo *complexus* – “o que é tecido junto”, e que vem recebendo tratamento de uma epistemologia própria, o Método do Pensamento Complexo, de Edgar Morin.

2.3 TEORIA DO PENSAMENTO COMPLEXO

Ao considerar que o complexo tem em sua gênese a mistura – que só existe enquanto produto e não como parcelas –, Edgar Morin rompe com o paradigma simplificador da ciência

¹³ No original, em inglês: “*The representation of knowledge in symbolic form is a matter that has preoccupied the world of documentation since its origin. [...] in all these fields it is necessary to decide how knowledge may be represented so that the representations may be manipulated.*”.

de tradição cartesiana que afirmava ser a clara distinção entre os objetos uma característica intrínseca da verdade. Deste modo, afirmar que a origem do conhecimento está na complexidade, significa retomar, em certa medida, a máxima aristotélica: que o todo precede as partes. Vera Dodebei recupera em “Continuidades e rupturas em representação do conhecimento” (DODEBEI, 2013) as notas de aula do professor Manoel Adolpho Wanderley (1980) e a dissertação defendida em 1973 na qual Wanderley apresenta restrições aos sistemas de classificação com base lógica, apresentando os mesmos argumentos:

Não é sempre exequível, ao dividir-se um gênero por características sucessivas, ater-se a uma regra de todo coerente [...] Contudo, nos sistemas tradicionais, a formação de classes pela combinação de características de todo distintas é inevitável, mormente quando se tratar de entidades complexas [...].

Há frequentes superposições quando se classificam e arranjam fisicamente, quanto ao conteúdo, os documentos, i. é, os objetos a que os termos das classes em última análise se referem. Verifica-se a impossibilidade de uma correspondência biunívoca entre documentos e termos de classes, o que se soma à natural complexidade das relações entre as próprias ciências e ao caráter interdisciplinar da pesquisa moderna (WANDERLEY apud DODEBEI, 2013, p.81. Grifos do autor).

Antes de aprofundar as medidas de abordagem a esta complexidade, é importante esclarecer a diferença entre os termos “complexo” e “complicado”, frequentemente tomados como sinônimos.

O filósofo Roberto Poli afirma que um problema complicado pode ser decomposto em partes e resolvido, como por exemplo, um complicado cálculo de integrais na Matemática. Já um problema complexo não pode ser resolvido pela decomposição e “uma vez que as totalidades indecomponíveis não são (inteiramente) compreensíveis pelas partes, a manipulação das partes pode gerar consequências inesperadas” (POLI, 2013, p. 144) ¹⁴.

Ao retrazar as origens do pensamento complexo na ciência ocidental, Edgar Morin situa a emergência da complexidade como uma resposta necessária à crise do conhecimento a partir do final do século XIX e agravada no século XX, especificamente ao que denomina como inequívoca constatação da insuficiência e falência do modelo científico de inspiração cartesiana frente aos avanços da Física e da Química quânticas, epitomizadas pela célebre frase de Albert Einstein que afirmou que o “materialismo morreu de asfixia por falta de matéria” ¹⁵.

¹⁴ No original, em inglês, “[...]since indecomposable wholes are not (entirely) understandable from their parts, manipulation of parts may engender unexpected consequences”. Tradução nossa.

¹⁵ Para a análise pormenorizada da emergência do paradigma complexo, ver a coletânea: MORIN, Edgar. *Introdução ao pensamento complexo*. 4. ed. Porto Alegre: Sulina, 2011.

Edgar Morin propõe um novo modelo de pensamento, que denomina *Pensamento Complexo*, para o qual desenvolve uma sistemática de análise e de operação, batizada de *O Método* ¹⁶, descrito em uma obra em seis volumes publicados entre 1977 e 2004, sendo o terceiro volume subintitulado *O Conhecimento do Conhecimento* ¹⁷ especialmente devotado ao tema do conhecimento, e cuja fundamentação conceitual e operacionalização metodológica servem diretamente a esta tese.

A teoria de Morin parte de uma premissa fundamental que afirma que “o pensamento complexo aspira ao conhecimento multidimensional. Mas ele [o pensamento complexo] sabe desde o começo que o conhecimento completo é impossível” (MORIN, 2005a, p. 6). A teoria, portanto, não almeja à completude e tem, no seu princípio operacional as medidas para considerar as incertezas e indeterminações próprias ao conhecimento humano. No mesmo trecho, Morin afirma que “não se trata de retomar a ambição do pensamento simplificador, que é a controlar e dominar o real. Trata-se de exercer um pensamento capaz de lidar com o real, de com ele dialogar e negociar” (ibidem, p.6).

Estas afirmações são compatíveis com as constatações apresentadas por Roberto Poli, quando este autor afirma que os sistemas complexos “nunca são totalmente apreensíveis por qualquer modelo existente, pois modelos de sistemas complexos – mesmo em princípio – são sempre incompletos e se tornam inconsistentes com o passar do tempo” (POLI, 2013, p. 145–146) ¹⁸.

Esta ruptura paradigmática proposta pelo Pensamento Complexo tem consequências radicais para a concepção científica. O modelo científico tradicional sempre rejeitou a incompletude e a imprecisão, tomando-as como falhas, como impossibilidades para a ciência. Retoma-se aqui novamente a discussão sobre a obrigatoriedade da integralidade e da inequivocidade no pensamento científico tradicional, o qual, especificamente nos estudos da Organização do Conhecimento, resulta na formulação das abordagens dicotômicas da classificação.

Romper com este paradigma não significa, como a ciência considerou por muitos séculos, pregar um “não modelo”, ou uma impossibilidade de se tratar a informação, mas sim, saber de antemão que o modelo produzido será parcial, com formulações que não são inequívocas nem mesmo em relação a um determinado ponto de vista (como já advertira

¹⁶ No original, em francês, *La Méthode*.

¹⁷ No original, em francês, *La Connaissance de la connaissance*.

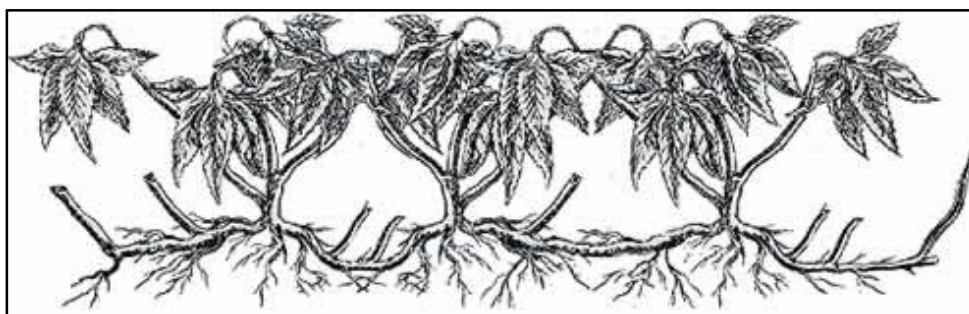
¹⁸ No original, em inglês, “complex systems [...] are never fully graspable by any model whatsoever: models of them - even in principle - are always incomplete and diverge over time”. Tradução nossa.

Langridge)), pois mesmo para este ponto de vista, o modelo nunca será completo nem imune às oscilações provenientes da própria natureza de disputas e transformações às quais qualquer sistema complexo está submetido (como advertiu Bourdieu). A este respeito, Roberto Poli sintetiza dizendo que a solução de “*aprender a dançar* com um sistema complexo é definitivamente diferente de *resolver* os problemas que surgem deste tipo de sistema” (ibidem, p.143) ¹⁹.

O proveito que se pode obter ao adotar um sistema complexo depende, portanto, diretamente de não se esperar deste tipo de sistema uma resposta definitiva e inequívoca, oferecida à custa de simplificações, como ocorre nos sistemas tradicionais.

Gilles Deleuze (1925-1995) e Félix Guattari (1930-1992) propuseram em “Mil Platôs” (DELEUZE; GUATTARI, 1995) o modelo de rizoma em oposição aos modelos estruturais como a raiz (a base que opera a constituição) e aos gerativos como a árvore (o pivô que instala a sucessão). Segundo os autores, o rizoma é o “único subtraído da multiplicidade a ser constituída” (ibidem, p.15).

Figura 5 – Representação gráfica de um rizoma



Fonte: (MURPHY; SMITH, 2001).

Na acepção deleuzeana, as fronteiras são zonas de trocas, que articulam um devir que é sempre duplo, e para cuja abordagem é proposta uma “hiper-hermenêutica” na qual o princípio não seja o de interpretar, mas de desdobrar e experimentar, a qual “mais que fixar-se no todo, produza a diferença” ²⁰ (DELEUZE; GUATTARI, 1980, p. 96). Em Deleuze, portanto, as fronteiras não são definidoras ou delimitadoras de “conceitos” tal como ocorre

¹⁹ No original, em inglês, “*learning to dance with a complex system is definitely different from ‘solving’ the problems arising from it*”. Tradução nossa. Grifos do autor.

²⁰ Tradução de Bento Prado Jr. e Alberto Alonso Munoz, na versão em português: “O que é a Filosofia?” Rio de Janeiro: Editora 34, 1992.

nos sistemas de organização do conhecimento tradicionais. Embora em Deleuze os “conceitos” também sejam produtos da diferenciação, estes devem ser concebidos “mais como abertura e atualidade do que como dado ou acabamento” (ibidem). Neste ponto, Morin e Deleuze, endossam a leitura de outro filósofo francês, Gaston Bachelard (1884-1962) quando este postula que uma ciência efetivamente nova precisa substituir “a clareza em si por uma espécie de luz operacional. Longe de ser o ser a ilustrar a relação, é a relação que ilumina o ser. [...] Na realidade, não existem fenômenos simples; o fenômeno é uma teia de relações.”

²¹ (BACHELARD, 1949, p. 144. Tradução nossa).

No intuito de alcançar esta visão, o sistema de organização do conhecimento proposto nesta tese realiza um deslocamento significativo, deixando de considerar a noção de conceito como a “unidade básica do conhecimento” (DAHLBERG, 1978) para tomar um “objeto complexo” como sendo esta unidade básica. Embora fosse possível denominar este “objeto complexo” também pelo termo “conceito”, parece pouco adequado, uma vez que há, de fato, uma alteração em sua intenção e também porque há um termo para designá-lo de maneira mais precisa: dispositivo.

2.4 DISPOSITIVO COMO UNIDADE BÁSICA DO CONHECIMENTO

A mudança de paradigma, anunciada por este trabalho e ilustrada pela metáfora imagética da passagem da “árvore” para o “rizoma”, se dá, portanto, através da ruptura com a noção de tradicional de organização do conhecimento a partir de conceitos. Esta passagem foi construída com o recurso a dois artefatos teóricos já conhecidos pela Organização do Conhecimento: o “dispositivo” e o “documento”.

A menção à noção de dispositivo consta na proposta de elaboração de um modelo mundial de sistema de informações sobre ciência e tecnologia, denominado UNISIST, realizado pela UNESCO e pela ICSU durante a década de 1970 (DAM CHRISTENSEN, 2014) ²².

A ideia de dispositivo, tal como adotada neste trabalho, remete à formulação apresentada por Michel Foucault (1926-1984). Em seus estudos sobre a noção de poder e suas

²¹ No original, em francês, “substitue à la clarté en soi une sorte de clarté opératoire. Loin que ce soit être qui illustre la relation, c'est la relation qui illumine l'être. [...] En réalité, il n'y a pas de phénomènes simples; le phénomène est un tissu de relations.”

²² Para uma revisita contemporânea ao modelo UNISIST pela ótica da análise de domínios, ver: HJØRLAND, B., SØNDERGAARD, T.F. ANDERSEN, J. UNISIST model and knowledge domains. In: *Encyclopedia of Library and Information Science*, [s.l.][s.n.], 2005, p.1-14.

relações, Foucault empregou o termo “dispositivos de poder” ²³ em “A Ordem do Discurso” (1971), desenvolvendo-a posteriormente em “Vigiar e Punir” (1975). Em entrevista concedida em 1977, Foucault proferiu o que parece ser a definição mais aceita do que vem a ser dispositivo:

[...] um conjunto decididamente heterogêneo que engloba discursos, instituições, organizações arquitetônicas, decisões regulamentares, leis, medidas administrativas, enunciados científicos, proposições filosóficas, morais, filantrópicas. Em suma, o dito e o não dito são os elementos do dispositivo. O dispositivo é a rede que se pode estabelecer entre estes elementos... [e entre estes] existe um tipo de jogo, ou seja, mudanças de posição, modificações de funções, que também podem ser muito diferentes, [cuja finalidade] é responder a uma urgência.” ²⁴ (FOUCAULT, 1977, [s.p.]).

A aproximação deste conceito com o abstrato “objeto complexo” enunciado por Morin se dá precisamente pela “teia de relações que iluminam o ser” apontada por Bachelard. Gilles Deleuze, em seu trabalho como comentador crítico de Foucault, retomou a noção de dispositivo, aproximando-a imgeticamente de um nó do rizoma, nó este formado por múltiplas linhas, nas quais é possível identificar as linhas de visibilidade e de enunciação e no qual há também as linhas de fissura (DELEUZE, 1988). Segundo Deleuze, é a partir deste “emaranhado” que o homem é capaz de vislumbrar como se constitui uma determinada racionalidade em uma determinada forma de organização. Para ele, “os dispositivos são máquinas de fazer ver e falar” (ibidem, p. 1) e tem como característica intrínseca, além de permitirem vislumbrar uma organização existente, atuarem como instrumento do duplo devir, pois para ele “todo o dispositivo se define, pois, pelo que detém em novidade e criatividade, o qual marca, ao mesmo tempo, sua capacidade de se transformar ou se fissurar em proveito de um dispositivo do futuro” (ibidem, p. 5-6) ²⁵.

Outra noção relevante para abordagem da complexidade é a relação entre “virtual” e “atual”. Segundo Deleuze, toda multiplicidade implica elementos “atuais” e “virtuais”, sendo o primeiro grupo complemento ou produto do segundo. O processo de atualização constitui a individualidade a partir da virtualidade. Esta relação ocorre no que Deleuze denomina “plano de imanência” que, conforme suas palavras “compreende a um só tempo o virtual e sua

²³ No original, em francês, “*les dispositifs de pouvoir*”.

²⁴ Do original, em francês, “*le dispositif, c’est justement la nature du lien qui peut exister entre ces éléments hétérogènes. Ainsi, tel discours peut apparaître tantôt comme programme d’une institution, tantôt au contraire comme un élément qui permet de justifier et de masquer une pratique qui, elle, reste muette, ou fonctionner comme réinterprétation seconde de cette pratique, lui donner accès à un champ nouveau de rationalité. Bref, entre ces éléments, discursifs ou non, il y a comme un jeu, des changements de position, des modifications de fonctions, qui peuvent, eux aussi, être très différents. Troisièmement, par dispositif, j’entends une sorte – disons – de formation, qui à un moment historique donné, a eu pour fonction majeure de répondre à une urgence.*”

²⁵ Tradução de Cláudia Sant’Anna Martins em DELEUZE, Gilles. *Foucault*. São Paulo: Brasiliense, 2005.

atualização, sem que possa haver aí limite assimilável entre os dois” (DELEUZE, 1996, p. 51). O sentido inverso ocorre no que Deleuze chama de “movimento de desdobramento de virtualidades” a partir do momento em que o atual coexiste com outros movimentos de aproximação do atual a partir de seu virtual. O dispositivo é o artefato teórico virtual capaz de “preservar a imagem atual do presente que passa e a imagem virtual do passado que se conserva distinguindo-se somente na atualização até se tornarem indiscerníveis, cada um se apropriando do papel do outro” (ibidem, p.55).

O filósofo italiano Giorgio Agamben retomou a noção foucaultiana de dispositivo expandindo a sua definição para “tudo aquilo que possui, de algum modo, a capacidade de capturar, orientar, determinar, interceptar, modelar, controlar e assegurar os gestos, as condutas, as opiniões e os discursos dos seres vivos” ²⁶ (AGAMBEN, 2007, p. 40. Tradução nossa).

Em uma vertente da Ciência da Informação que vem sendo desenvolvida por pesquisadores brasileiros e franceses (e da interação entre eles), Regina Marteleto e Viviane Couzinet empregam a noção de dispositivo para designar o construto que viabiliza o estudo analítico das mediações em redes, denominando-o “dispositivo de informação e comunicação”, ou mais sucintamente “dispositivo infocomunicacional” (MARTELETO; COUZINET, 2013). É precisamente nesta acepção que o presente trabalho adota o termo dispositivo, proposta como “agenciamento de elementos funcionando conjuntamente, [que] exprime a sensibilidade ao elo e a necessidade de um projeto cujo resultado, perenidade e força só podem existir no coletivo” (ibidem, p. 8).

Portanto, um dispositivo infocomunicacional é um elemento que participa em um projeto maior, dentro do qual aflui para uma finalidade específica. Gonzalez de Gómez destaca que um dispositivo só se define dentro de seu campo de operação, uma vez que “não possui a neutralidade que o torna um instrumento para qualquer orientação ou intenção posterior”, mas que possui, contudo, “desde seu início, regras de formação e transformação” (GONZÁLEZ DE GÓMEZ, 1999, p. 26). Esta visão é coincidente com o princípio operacional proposto por Michel Foucault, em “Arqueologia do Saber” (1969), na qual o autor francês afirma que são as relações entre os elementos constitutivos dos objetos, aparentemente dispersos no discurso, que permitem que sejam identificadas as regularidades e

²⁶ Do original, em francês “*tout ce qui a, d’une manière ou une autre, la capacité de capturer, d’orienter, de déterminer, d’intercepter, de modeler, de contrôler et d’assurer les gestes, les conduites, les opinions et les discours des êtres vivants*”.

as recorrências. Essas relações apresentam-se em um sistema de dependências que obedece a regras e que pode ser organizado (FOUCAULT, 1969).

São, portanto, as relações estabelecidas, na medida em que formam regularidades em meio à aparente dispersão, remetendo a um sistema de regras de formação e transformação que regem a formação de uma unidade ou objeto complexo. Neste sentido, o dispositivo infocomunicacional possui uma característica de virtualidade em relação ao objeto complexo, que precisa ser atualizado na forma de um documento.

A retomada da centralidade da noção de documento vem sendo discutida na Organização do Conhecimento contemporânea, a partir da constatação que ele “revela o caráter social e simbólico da informação e, conseqüentemente, os ambientes e as situações concretas de uso” (ORTEGA; LARA, 2010). O documento é, portanto, um objeto concreto no qual o conteúdo ganha forma e é, ao mesmo tempo, o suporte que possibilita a sua circulação (COUZINET, 2009).

A noção de documento, familiar à Organização do Conhecimento, remete inevitavelmente aos trabalhos fundadores de Paul Otlet (1868-1944) e Henri La Fontaine (1854-1943) e aos desenvolvimentos posteriores conduzidos por Suzanne Briet (1894-1989) (BUCKLAND, 2007). Embora a distinção entre conteúdo e suporte houvesse sido formulada por Otlet no “Tratado da documentação” (1934), foi a partir da releitura realizada por Jean Meyriat (1921-2010) que passou a ser estudada a “dimensão comunicacional” do documento (MEYRIAT, 1981). Neste sentido, Marteleto e Couzinet justificam a retomada da noção de documento como meio de “alcançar a variedade e a complexidade de suas materialidades, apropriações e usos nas práticas sociais” (MARTELETO; COUZINET, 2013, p. 7).

Considerando a canção popular brasileira como o domínio escolhido para exemplificar e validar os artefatos teóricos propostos nesta tese, toma-se a canção – virtualmente qualquer canção – como dispositivo infocomunicacional. O documento através do qual se atualiza este objeto complexo e que permite acessar a sua rede de relações e abordar a sua complexidade é o seu registro material mais imediato, a gravação.

3 OBJETO DE PESQUISA: CANÇÃO POPULAR BRASILEIRA

Para os americanos
 Branco é branco, preto é preto
 E a mulata não é a tal
 Bicha é bicha, macho é macho
 Mulher é mulher, e dinheiro é dinheiro
 E assim ganham-se, barganham-se, perdem-se
 Concedem-se, conquistam-se direitos
 Enquanto aqui embaixo a indefinição é o regime
 E dançamos com uma graça cujo segredo nem eu mesmo sei
 Entre a delícia e a desgraça
 Entre o monstruoso e o sublime (VELOSO, 1992).

O trecho acima é da letra da canção “Americanos” de Caetano Veloso, gravada em 1992 no álbum “Circuladô Vivo”, na qual o Veloso tece uma visão dos americanos do norte – mais precisamente dos estadunidenses – refletida a partir do espelho dos americanos do sul – especificamente, dos brasileiros ²⁷. A distinção marcada pelo artista identifica nos vizinhos do norte uma visão pragmática (no sentido que uma ideia corresponde diretamente aos seus desdobramentos práticos) enquanto “aqui embaixo” há uma permanente oscilação, simbolizada na figura da “dança graciosa”, que representa a indefinição que nos rege. Se para os estadunidenses as coisas do mundo são classificáveis de maneira inexorável (“branco é branco”, “preto é preto”), aqui embaixo há um “estar entre” as categorias, multiplicando-as em vez de reduzi-las, e ampliando o alcance dos sentidos em vez de encurtá-lo.

A canção de Caetano Veloso ilustra que a “mulata” – representante prototípica da miscigenação uma vez que não é branca nem é preta e sim resultado da mistura entre ambos – não tem lugar neste tipo de classificação (“E a mulata não é a tal”).

José Miguel Wisnik, polímata estudioso da canção popular brasileira destaca o “singular aspecto cultural formativo” que a canção popular assume no Brasil ao permitir tecer “uma rede de recados e diálogos na vida cultural do país: a formação de um protoespaço público, configurando uma espécie de repertório comum de experiências sociais e estéticas ao transportar conhecimentos e linguagens entre as diferentes classes e circuitos culturais” e, declara que a canção popular constitui-se, como “uma nova forma de ‘Gaia Ciência’, isto é, um saber poético-musical” peculiar e específico da cultura brasileira (WISNIK, 2004, p. 218).

²⁷ Este exemplo foi apresentado anteriormente pelo autor em trabalho escrito com Rosali Fernandez de Souza (DE SANTIS; SOUZA, 2013). A sua utilização nesta tese, contudo, resulta de uma nova leitura, revista e expandida.

3.1 CANÇÃO POPULAR ENQUANTO DOMÍNIO AUTÔNOMO

A tarefa de abordar as canções populares brasileiras começa com a necessidade de se determinar o que é este domínio de estudo e identificar as fronteiras (sempre fluidas e sobrepostas) que o situam em relação a outros domínios artísticos, como os já citados domínios da Música e da Poesia e outros como a Literatura, o Cinema, a Política, a Economia etc.

Definir “o que é canção popular” é mais difícil do que pode parece à primeira vista. Embora seja amplamente difundida e consumida como produto cultural na sociedade contemporânea, foi somente a partir da década de 1990 que a canção popular passou a ser considerada de forma sistemática um objeto de estudo pelo campo científico. Arthur Nistrovski considera que pode se atribuir a esta falta de reflexão científica a respeito da canção popular boa parte da sua imprecisão teórica (NESTROVSKI, 2007).

Em estudos na área de Musicologia é comum considerar a canção popular como um subproduto, um desvio da canção clássica. Nestes casos, a canção popular é usualmente tomada como uma música mal-feita e da qual eventualmente interessa extrair um ou outro aspecto melódico ou rítmico para estudos musicológicos. Já em abordagens de estudos Linguísticos e Sociológicos, a visão ainda corrente entende a canção popular como sendo um tipo de música feito a partir de um texto literário e, neste caso, usualmente aparta-se a música para, a partir dos textos - as letras das canções - fazer análises de natureza textual. Uma terceira acepção é encontrada em trabalhos que utilizam a canção popular como exemplo, embora não se detenham em estudá-la especificamente - como ocorre muitas vezes na Ciência da Informação e na Biblioteconomia. Nestes casos, a canção popular é usualmente tomada como sinônimo de música popular, às vezes até de música em geral.

Em um dos raros trabalhos anteriores à década de 1990 que menciona o assunto (embora seja fundamentalmente dedicado à canção clássica ²⁸), Vasco Mariz, especialista em musicologia histórica, revela que a origem do conceito de canção:

[...] recua até a mais remota antiguidade e teria começado por uma forma especial de onomatopeia. Quando a civilização grega engatinhava ainda e não existia a escrita, cantava-se em honra dos deuses [...]. Das melodias litúrgicas parecem haver brotado as canções populares mais primitivas. (MARIZ, 1977, p. 26).

²⁸ Nesta tese opta-se por empregar o termo “clássica” em lugar de “erudita”, quer para a canção, quer para a música como um todo. Ambos os termos são controvertidos e possivelmente inadequados, considerando-se, portanto, não haver prejuízo para o leitor nesta escolha.

Desta afirmação, surgem assunções importantes, como a identificação que a origem da canção é a canção popular e não a canção clássica e também que a canção surge espontaneamente antes mesmo do surgimento da escrita. Estudos recentes, aparentemente tomando outros caminhos, chegaram à mesma conclusão: a canção popular é um tipo de objeto específico, não só pela sua anterioridade em relação a outras formas musicais ou literárias, mas principalmente por apresentar uma linguagem própria (TATIT, 1994).

A canção popular nasce da entoação da voz. A mesma voz que fala pode ser usada para entoar e, desta entoação, nasce a música da canção popular. O texto – o que é dito pela voz – nasce junto, misturado a essa música. E essa mistura entre voz e canto – entre letra e música – oferece infinitas possibilidades de organização, gerando a instabilidade que é típica das canções populares e que é levada ao limite pela canção popular brasileira.

3.2 CANÇÃO POPULAR ENQUANTO DOMÍNIO COMPLEXO

Branca é branca, preta é preta
Mas a mulata é a tal, é a tal!
[...]
Ai mulata, cor de canela!
Salve, salve, salve, salve, salve ela! (JOÃO DE BARRO; ALMEIDA, 1947).

Este trecho é da canção “A mulata é a tal”, de autoria de João de Barro e Antônio Almeida, gravada em 1947 por Ruy Rey²⁹ e que foi sucesso no carnaval de 1948. É razoável presumir que a canção “Americanos” de Caetano Veloso alude e dialoga com ela.

Na letra de “Americanos”, Caetano Veloso usa uma oração coordenada aditiva para negar a superioridade da mulata: “*E a mulata não é a tal*”. A frase é peremptória: lá a mulata não tem vez. Ao contrário, na letra de “A mulata é a tal”, a afirmação de superioridade da mulata é feita através de uma oração coordenada adversativa. “*Mas a mulata é a tal*”, que indica a inclusão e o reconhecimento da branca e da preta, embora o autor prefira a mulata. Esta diferença linguística, que determina e marca a diferença de sentido do discurso nas duas canções, é a síntese das visões “lá de cima” e “daqui de baixo”, ou seja: rejeitar ou assimilar.

Tornada “natural” na canção do final da década de 1940 e contraposta à cultura estadunidense no início da década de 1990 (nos dois excertos musicais apresentados), essa

²⁹ João de Barro é o pseudônimo do compositor Carlos Alberto Ferreira Braga (1907-2006), também conhecido no meio artístico como Braguinha.

Antonio Almeida (1911-1985) foi um compositor popular brasileiro.

Ruy Rey é o nome artístico de Domingos Zeminian (1915-1995), cantor, compositor e ator.

cultura brasileira mestiça foi socialmente construída a partir do esforço de definição de uma “cultural nacional popular”, promovida pelo governo de Getúlio Vargas no início da década de 1930. Historicamente, as elites brasileiras renegavam a cultura popular, especialmente aquela de origem negra. Este processo de elevar a cultura mestiça ao lugar mais alto da identidade nacional promovido pelo projeto nacionalista de Vargas foi diretamente influenciado pelo pensamento de intelectuais como Gilberto Freyre, considerado o fundador do “mito da democracia racial no Brasil” (FREYRE, 1933). Atualmente, no entanto, é reconhecido que a alteridade religiosa e estética das culturas negras e brancas no Brasil não é nem nunca foi pacífica. Esta “harmonia” foi forjada a partir de iniciativas governamentais que podem ser percebidas através da cultura, com o emblemático momento em que os negros passaram a ocupar os espaços públicos com a sua música e dança ao mesmo tempo em que os brancos apadrinhavam o samba e o choro e passavam a frequentar os terreiros de candomblé. Este percurso a partir dos encontros de Villa-Lobos e Gilberto Freyre (brancos representantes da cultura da elite) com Pixinguinha, Donga e Patrício Teixeira³⁰ (músicos negros de origem pobre), são descritos no livro “O Mistério do Samba”, de Hermano Vianna, e reconstroem as origens de um fenômeno possivelmente sem precedentes em outras partes do mundo, esta assimilação e transformação da cultura popular marginal em cultura oficial (VIANNA, 1995).

Ao compatibilizar a “sinfonização nacionalista” encarnada pelo maestro e compositor Villa-Lobos em favor da “estratégia de controle simbólico da totalidade social” com a legitimação e a apropriação do carnaval, que de “movimento ofensivo da estratégia de afirmação dos grupos marginais” passou a ser expressão oficial da espontaneidade da cultura brasileira (WISNIK, 1982, p. 160–161), Vargas aproveitou-se da mestiçagem latente na cultura popular para atenuar a luta de classes e criar uma aliança em torno de seu ideal – populista – de governo (SEVERIANO, 1983).

A latência da cultura popular mestiça retratada por Freyre se manifesta de forma exemplar na canção brasileira. Até o século XIX, a música popular era cantada nos bairros

³⁰ Getúlio Dornelles Vargas (1882-1954), advogado e político brasileiro, foi presidente do Brasil em dois períodos: de 1930 a 1945, e de 1951 a 1954.

Gilberto de Mello Freyre (1900-1987), polímata brasileiro (sociólogo, antropólogo, historiador, escritor e pintor).

Heitor Villa-Lobos (1887-1959), maestro e compositor brasileiro.

Pixinguinha, apelido de Alfredo da Rocha Vianna Filho (1897-1973), maestro, flautista, saxofonista, compositor e arranjador brasileiro.

Donga, apelido de Ernesto Joaquim Maria dos Santos (1889-1974), músico, compositor e violinista brasileiro.

Patrício Teixeira(1893-1972), cantor e violonista brasileiro.

afastados e nos becos e ruelas por negros e mestiços em estilos como o lundu, a modinha e o maxixe, alcançando fama restrita aos seus meios informais de circulação (TINHORÃO, 1998). A obra de artistas como Joaquim Manoel da Câmara (1780?-1840?) e José Francisco Leal (1792-1829), só chegaram a nossos dias graças à intervenção de artistas estrangeiros, como é o caso de Sigismund Neukomm (1778-1858), compositor e pianista austríaco, integrante da Missão Artística Francesa no Brasil, que coletou e registrou a música popular que era cantada no Rio de Janeiro, levando-a para a Europa (LANZELOTTE, 2009).

Na virada para o século XX, Ernesto Nazareth (1863-1934) constrangia ainda mais as fronteiras entre o clássico e o popular, compondo acachapantes sucessos populares em obras de elevada qualidade musical (ANDRADE, 1962), (WISNIK, 2008). Com Ernesto Nazareth, Chiquinha Gonzaga (1847-1935) e Catulo da Paixão Cearense (1863-1946), a música popular deixava de se restringir às ruas e passou a frequentar, ainda que de forma tímida, os salões da alta sociedade e o palácio do então presidente Hermes da Fonseca (VIANNA, 1995). Os Oito Batutas³¹ tocavam para a elite carioca, sendo inclusive financiados pelo aristocrata Arnaldo Guinle (1884-1963) para excursionarem em Paris durante seis meses no ano de 1922 (ibidem).

Esta mistura entre a cultura clássica e a cultura popular em estado latente foi então tornada projeto político do país a partir da década de 1930, proporcionando o surgimento de formas ainda mais insuspeitadas de criação. Dentro deste contexto, a canção popular, que tem na sua raiz a mistura entre a voz e o canto, passa a misturar uma variedade de vozes e de cantos. A miríade de temas, assuntos, estilos, formas de se compor e de se interpretar canção popular no Brasil não encontra paralelo em outras partes do mundo e tem como epítome o compositor, cantor, poeta e diplomata Vinícius de Moraes (1913-1980). Premiado poeta, diplomata de carreira e autor de mais de 300 canções populares, Vinícius se declarava “o branco mais preto do Brasil” e transitava com naturalidade entre esses diversos ofícios e a multiplicidade de formas e estilos da canção popular.

Embora seja um caso notável (por sua prolificidade e qualidade), Vinícius de Moraes não foi uma exceção. Músicos brasileiros de formação clássica como Radamés Gnattali (1906-1988), César Guerra-Peixe (1914-1993) e Tom Jobim (1927-1994), envolveram-se diretamente no fazer da música popular, influenciando e sendo influenciados pelos músicos

³¹ Conjunto musical brasileiro formado por Pixinguinha na flauta, Donga e Raul Palmieri no violão, Nelson Alves no cavaquinho, China no canto, violão e piano, José Alves no bandolim e ganzá e Luis de Oliveira na bandola e reco-reco. Verbete “Oito Batutas” Dicionário Cravo Albin da Música Popular Brasileira. Online. Disponível em: <<http://www.dicionariompb.com.br>>. Acesso em: maio 2016.

populares, fenômeno que ocorre até os dias atuais com músicos como Turíbio Santos, Egberto Gismonti, Gilson Peranzetta e Mauro Senise, para citar apenas alguns exemplos. As canções de compositores pouco escolarizados como Cartola e Wilson Batista ³², a técnica vocal de Orlando Silva ³³, ou as invenções harmônicas de Jorge Ben Jor são exemplos desta espontânea capacidade criativa e de elevada qualidade musical (melódica, rítmica, harmônica, lírica e poética) que se manifesta amiúde na música popular brasileira.

Os cancionistas – termo empregado por Luiz Tatit para se referir aos que se ocupam deste ofício singular de fazer canções – tem em suas obras essa mistura e essa pluralidade que os torna refratários a categorizações absolutas (TATIT, 1995), articulando temas que vão de uma ida à praia até reflexões metafísicas, oscilando entre a linguagem coloquial e a culta, entre formas musicais primitivas e as mais sofisticadas ³⁴, colocando-se indefinidamente “entre a delícia e a desgraça / entre o monstruoso e o sublime”, como apontam os versos da canção de Caetano Veloso.

3.3 CANÇÃO POPULAR ENQUANTO DISPOSITIVO INFOCOMUNICACIONAL

*And I told about equality
And it's true
Either you're wrong or you're right
[...]
But if you're thinkin' about my baby
It don't matter if you're black or white* (JACKSON, 1991).

A visão dos americanos do norte anunciada por Caetano Veloso, na qual “branco é branco”, “preto é preto” permite explorar outras relações. Uma delas é com a canção *Black or White* do artista estadunidense Michael Jackson (1958-2009), sucesso mundial em 1991 (ano anterior ao lançamento de *Americanos*). A própria biografia de Michael (um negro que buscou ao longo dos anos embranquecer a sua pele) permitiria identificar os elementos que transbordam em sua canção: ao pregar a igualdade, o artista recorre a dicotomias: “ou se está certo, ou se está errado”; e diz que não importa se a pessoa “é preta” ou “é branca”. Nos dois casos, a separação dicotômica precede e prevalece.

³² Cartola, apelido de Angenor de Oliveira (1908-1980), compositor e cantor.

Wilson Baptista de Oliveira (1913-1968) compositor.

³³ Orlando Garcia da Silva (1915-1978), intérprete de canções populares que, intuitivamente revolucionou a maneira de cantar música popular.

³⁴ Concretizam este exemplo, respectivamente, “Atoladinha”, de MC Bola de Fogo e “Bolero Blues” de Chico Buarque e Jorge Helder.

A habilidade do cancionista brasileiro de construir a complexidade tecendo uma rede de relações imanentes que são atualizadas na forma de documento através da gravação permite “vislumbrar o emaranhado” na acepção deleuzeana, ou, abordar o complexo, conforme Morin. As possibilidades informacionais e comunicacionais do dispositivo canção popular são potencialmente infinitas. Puxando novos fios do emaranhado, é possível, por exemplo, relacionar a canção *Americanos* com outra canção de Caetano Veloso, *Sugar Cane Fields Forever*, lançada em 1973, na qual o cancionista declara em primeira pessoa:

Sou um mulato nato
No sentido lato
Mulato democrático do litoral (VELOSO, 1973)

O título, por sua vez, alude inevitavelmente a outro sucesso internacional, *Strawberry Fields Forever*, dos Beatles (1967), e é creditado como uma parceria entre Caetano e o poeta maranhense Sousândrade³⁵. O caráter experimental, o rigor estético e o cunho revolucionário do cancionista e do poeta permitem identificar novas relações no plano social e político. Sousândrade era republicano militante durante o Império, viveu e escreveu grande parte da sua obra em Nova Iorque, nos Estados Unidos, retornando ao Brasil para aderir ao movimento de proclamação da república em 1889 (CAMPOS; CAMPOS, 1982). Caetano, engajado na luta pela democracia, foi preso durante a ditadura militar e exilado em Londres entre 1969 e 1971. Ao aludir no título a um lugar que será “para sempre campos de cana de açúcar” e ao se autodeclarar como proveniente do litoral, é possível supor que Caetano refere-se a Santo Amaro, sua terra natal, outrora próspera com a indústria canavieira.

Ao abordar intuitivamente a rede de relações a partir de um dispositivo infocomunicacional, no caso a canção *Americanos*, há uma dupla intenção: primeiro a de exemplificar a sua aplicabilidade, verificando a sua pertinência em relação ao referencial teórico enunciado; segundo a de evidenciar a necessidade de ferramentas metodológicas adequadas para lidar com o domínio complexo.

Estas ferramentas devem considerar a característica de sobreposição e de fluidez das fronteiras estabelecidas pelos dispositivos infocomunicacionais ao mesmo tempo em que permitam a construção de um sistema de organização do conhecimento capaz de manipular os documentos relativos a estes dispositivos. Esta metodologia precisa considerar, portanto, a asserção de Roberto Poli, quando afirma que para lidar com a complexidade é preciso

³⁵ Joaquim de Sousa Andrade (1833-1902).

“aprender a dançar com ela”. Caetano Veloso, mais de duas décadas antes do filósofo italiano indicava que o cancionista já havia resolvido a questão: “E dançamos com uma graça cujo segredo nem eu mesmo sei”.

4 ASPECTOS METODOLÓGICOS RELACIONADOS À ORGANIZAÇÃO DO CONHECIMENTO

A consecução de um sistema de organização do conhecimento capaz de abordar domínios complexos e manipular dispositivos em lugar de conceitos é alcançada a partir de um enfoque metodológico heterodoxo e heterogêneo, incluindo:

- O *Método Complexo*, definido por Edgar Morin como instrumento da sua teoria do Pensamento Complexo;
- O esquematismo tensivo, a partir da formulação de estrutura e de “termo complexo” segundo os trabalhos originais de Viggo Brøndal e dos desenvolvimentos contemporâneos liderados por Claude Zilberbeg;
- A lógica fuzzy, a partir dos princípios básicos de funcionamento propostos por Lotfi Zadeh;
- O ambiente da web e sua arquitetura rizomática e centrada em documentos, a partir do seu criador, Tim Berners-Lee;
- A Organização do Conhecimento orientada pela “recuperação”, como alternativa aos sistemas tradicionais orientados pela “representação”. Esta formulação segue as formulações defendidas por Jesse Shera e Brian Vickery, atualizadas pelas tendências atuais da web.
- O aparato computacional, com o intuito de evidenciar tanto a sua base de construção – que impõe limites definitivos para os sistemas de organização do conhecimento –, quanto suas potencialidades e desenvolvimentos recentes, especialmente a partir das noções de web semântica e a utilização da linguagem OWL e de propriedades fuzzy.

4.1 MÉTODO COMPLEXO

Para a superação do modelo simplificador, Edgar Morin propõe como premissa do Pensamento Complexo a já mencionada “fluidez das fronteiras” e a sua possibilidade de mutua interferência. Para operacionalizar este modelo, Morin estabelece que a abordagem ao complexo deve partir do centro (e não de suas fronteiras) e define três princípios operativos gerais: o princípio dialógico, o princípio da recursividade e o princípio hologramático.

Estes três princípios são aqui apresentados levando em consideração a noção de dispositivo infocomunicacional anteriormente discutida. Embora o termo “dispositivo” não

seja empregado por Morin em “O Conhecimento do Conhecimento”, considerou-se a compatibilidade teórica das formulações, especialmente pelo frequente recurso de Morin à obra de Deleuze e ainda porque a formulação deleuzeana de dispositivo (DELEUZE, 1988) é posterior à publicação do volume 3 do “Método” (original de 1986).

O princípio dialógico é definido “como a associação complexa de instâncias necessárias em conjunto à existência, ao funcionamento e ao desenvolvimento de um fenômeno organizado (MORIN, 2012, p. 110. Grifos do autor). Na aplicação ao sistema de organização do conhecimento proposto nesta tese, as instâncias são os documentos. Um documento é o fenômeno “organizável” a partir do qual é possível abordar o conjunto de associações complexas de um dispositivo. O princípio dialógico, portanto, permite representar a dualidade na essência da unidade. De acordo com Morin, as relações complementares, concorrentes e antagônicas entre documentos constituem, na origem, um mesmo fenômeno (ibidem).

O segundo princípio do Pensamento Complexo é o da recursão organizacional, que propõe a ruptura com a ideia linear de causa/efeito, produto/produtor. Este princípio afirma que os efeitos retroagem sobre as causas, e que tudo o que é produzido volta-se sobre o que produz na forma de um circuito. Porém, mais que retroação, a recursão organizacional “trata-se de um processo em que os efeitos ou produtos são, ao mesmo tempo, causadores e produtores no próprio processo, sendo os estados finais necessários à geração dos estados iniciais” (ibidem, p.113). Este princípio é compatível com o processo de atualização de um dispositivo infocomunicacional, uma vez que, sendo as relações simultaneamente causas e causadoras, um dispositivo torna-se “ele mesmo autoconstrutivo, auto-organizador e autoprodutor” (ibidem, p.113). Se para o pensamento humano, a recursão é um desafio, pois escapa à concepção linear de obtenção dos resultados a partir das premissas, no processamento computacional, entretanto, é um recurso básico, corriqueiramente alcançado através de funções que operam sobre estruturas de pilhas de dados, controladas pela própria memória computacional ³⁶.

O terceiro princípio do Pensamento Complexo é o princípio hologramático, que, como o nome sugere, advém do conceito de holograma, técnica desenvolvida por Dennis Gabor (1900-1979) em 1948 e que permite o registro de padrões de interferência de luz, gerando

³⁶ Para detalhes sobre operações recursivas em computadores, ver: ROGERS JR., Hartle. *Theory of recursive functions and effective computability*. New York: McGraw-Hill, 1967.

representações em três dimensões com uma característica única: em um holograma, cada parte possui a informação do todo.

Este princípio não apenas confirma a máxima aristotélica anteriormente enunciada que o todo antecede as partes, mas a expande, ao permitir afirmar que “não apenas a parte está no todo, mas o todo está na parte”. (ibidem, p.114).

Este princípio está presente também na concepção da filosofia das formas simbólicas de Cassirer, que explicitamente afirma:

O todo e suas partes estão entrelaçados um no outro, como que ligados um ao outro pelo destino – e assim permanecem, mesmo quando se separaram um do outro puramente de fato. Aquilo que, depois dessa separação, é infligido à parte, também é infligido ao todo. Quem se apodera de uma mínima parte do corpo de uma pessoa, até mesmo de seu nome, de sua sombra, de sua imagem em espelho – que no sentido da intuição mítica, são igualmente ‘partes’ reais dele – com isso tomou posse dela, conseguiu um poder mágico sobre ela (CASSIRER, 2001b, p. 100).

No que diz respeito à construção de sistemas complexos de organização do conhecimento, este princípio está diretamente relacionado ao reconhecimento da relativa autonomia das partes – o que possibilita a atualização dos dispositivos e sua identificação a partir dos documentos –, mas permite afirmar que cada dispositivo dispõe de aspectos gerais compartilhados a partir dos quais é possível estabelecer as regras que permitem abordar o todo.

Estes três princípios do Pensamento Complexo operam em conjunto e constituem a base teórica de um método geral para abordagem da complexidade. No entanto, a proposta de elaboração de um sistema de organização do conhecimento pressupõe que seja estabelecida uma estrutura, a qual permitirá operar os dispositivos de maneira estável, ou seja, sistemática. Esta estrutura, no entanto, não deve recair na problemática da simplificação das definições e operações, devendo ser uma estrutura capaz de operar as instabilidades da complexidade. Para esta tarefa, são adotados princípios do esquematismo tensivo.

4.2 ESQUEMATISMO TENSIVO

Para enfatizar a característica inovadora da noção de tensividade é importante apresentar, mesmo que resumidamente, o seu principal antecedente: o estruturalismo linguístico, que tem em Ferdinand de Saussure (1857-1913) seu maior representante.

Saussure é apontado como o responsável pela consolidação dos estudos da linguística moderna que estabeleceram as bases do estruturalismo no século XX. Ao propor as dicotomias linguísticas (língua vs. fala; sincronia vs. diacronia; sintagma vs. paradigma e

significante vs. significado), Saussure elaborou uma teoria do valor, na qual postula que os signos linguísticos estão numa relação de oposição dentro do sistema de uma língua. Isto significa que um signo só adquire valor na medida em que não é outro signo qualquer, ou seja, um signo só *é* aquilo que os outros signos *não são* ³⁷. Os limites desta teoria foram investigados por Bourdieu em “O que falar quer dizer: a economia das trocas linguísticas” (BOURDIEU, 1982) e evidenciam a sua incompatibilidade com qualquer abordagem de cunho social, com as noções de fluidez, sobreposição e gradação. Entretanto, foi precisamente a partir da Linguística Estrutural e dos estudos desenvolvidos para superá-la que consiste a tensividade, que prega “o apego à estrutura, mais que ao estruturalismo” ³⁸ (ZILBERBERG, 2006a, p. 15. Tradução nossa).

Um dos precursores dos estudos linguísticos estruturais e não dicotômicos, foi o dinamarquês Viggo Brøndal, que por ser o mais desconhecido dos autores que compõem a base teórica desta tese, contará com um breve perfil biográfico, que visa a auxiliar a elucidação do seu papel decisivo nos estudos da tensividade e de sua influência na proposta aqui desenvolvida.

Rasmus Viggo Brøndal nasceu em 1887 em Copenhague e morreu na mesma cidade em 1942. Foi, portanto, da geração seguinte a de Saussure, e trabalhou como filólogo e professor de literatura sempre na Universidade de Copenhague. Pouco valorizado pelos seus pares, no seu necrológio, Jen Holt fez constar a irônica pergunta: se Brøndal fora um linguista dublê de filósofo ou um filósofo dublê de linguista (HOLT, 1943).

No centenário de seu nascimento, a renomada revista francesa *Langages*, dedicou-lhe uma edição especial em cujo prólogo Svend Erik Larsen retomou a pergunta de Holt atualizando-a para questionar quem celebraria aquela data: se a Linguística ou a Filosofia, e concluía que nem uma área nem outra, pois a visão estruturalista de Brøndal a respeito da linguagem ainda desagradava aos filósofos, e a sua defesa da primazia da lógica sobre a linguística o tornava mal visto pelos linguistas (LARSEN, 1987). Larsen, no entanto, defende que a obra de Brøndal estava à frente do seu tempo, e que precisa ser relida e (re)considerada, especialmente pelo fato de as abordagens inter e multidisciplinares deixarem de ser vistas como problemáticas no âmbito das discussões científicas:

Os domínios científicos que se beneficiarão do trabalho Brøndal serão aqueles que consideram que o confronto entre várias ciências não é um problema, mas um

³⁷ Para uma leitura completa, ver SAUSSURE, Ferdinand de. Curso de Linguística Geral. São Paulo: Cultrix, 1972.

³⁸ No original, em francês: “*l’attachement à la structure plutôt qu’au structuralisme*”.

desafio necessário para o desenvolvimento do pensamento científico. Serão aqueles que entenderem ser essencial proceder a uma reflexão paralela sobre os problemas metodológicos e epistemológicos de conceitos fundamentais básicos. Serão aqueles que veem na linguagem e na produção de sentidos sua área central da investigação e não apenas considerem estes fenômenos em sua imanência. As ciências humanas que apresentarem tal atitude - seja ela a Linguística, a Filosofia ou a Semiótica - serão aquelas que encontrarão a maior inspiração nas obras de Viggo Brøndal. (ibidem, p.7. Tradução nossa)³⁹.

Resgatada inicialmente pelo destacado linguista Algirdas Julien Greimas (1917-1992), a obra de Brøndal apresenta ainda hoje um desafio para sua apreensão, que é a introdução presumivelmente original em seu tempo, da noção de “complexidade” dentro da concepção de estrutura. O pesquisador Claude Zilberberg, estudioso contemporâneo da tensividade, afirma que a teoria elaborada por Brøndal “se caracteriza por uma dupla ambição: visa não apenas a descrever as diversas possibilidades de organização [...], mas também os possíveis devires de um sistema” (ZILBERBERG, 1987, p. 60)⁴⁰.

Brøndal afirma que um sistema deve ser simultaneamente dinâmico e estável. E que os termos são definidos na interseção entre essas duas dimensões, com a parte dinâmica (o sensível), prevalecendo sobre a estrutura (o inteligível) (ZILBERBERG, 2006b, p. 81). Como visto anteriormente, esta concepção é próxima à de Cassirer, e é reafirmada quando este último assevera que “não fosse por um sentido expressivo que se manifesta em certas vivências perceptivas, a existência permaneceria muda para nós” (CASSIRER, 2001b, p. 90). Sendo intencionalmente repetitivo, convem lembrar a máxima deleuzeana: “os dispositivos são máquinas de fazer ver e falar”.

As coincidências (no sentido de justaposições e não de acasos) são ainda mais significativas quando confrontadas à teoria do Pensamento Complexo. Nas palavras do próprio Brøndal, explicando a noção de estrutura:

“Concebemos aqui a estrutura como objeto autônomo e, por conseguinte, não derivável dos elementos, dos quais ela não é nem um agregado, nem uma soma; é por isso que se deve considerar como da maior importância o estudo dos sistemas

³⁹ No original, em francês: “*Les domaines scientifiques qui pourront tirer profit des travaux de Brøndal seront ceux qui considèrent que la confrontation entre plusieurs sciences n'est pas un problème, mais un défi nécessaire pour le développement de la pensée scientifique. Ce seront ceux qui jugeront qu'il est indispensable de mener une réflexion parallèle sur les problèmes de méthode et la base épistémologique des concepts fondamentaux. Ce seront ceux qui verront dans le langage et la production de sens leur champ de recherche central et qui ne se contentent pas de considérer ces phénomènes dans leur immanence. Les sciences humaines qui représentent une telle attitude — qu'elles s'appellent linguistique, philosophie ou sémiotique — seront celles qui trouveront le plus d'inspiration dans les œuvres de Viggo Brøndal.*”

⁴⁰ No original, em francês: “*La conception du linguiste danois se caractérise par une double ambition: elle vise non seulement à décrire les diverses possibilités d'organisation [...], mais encore les devenirs possibles d'un système.*” Tradução cotejada com a de Ivã Carlos Lopes, Luiz Tatit e Waldir Beividas, em Razão e Poética do Sentido (ZILBERBERG, 2006b).

possíveis e de sua forma. – E, no entanto, não poderíamos considerar os elementos que fazem parte de um sistema como simples derivados das correlações ou oposições estruturais; com efeito, importará distinguir entre as propriedades puramente formais de um sistema e sua matéria ou substância, que, mesmo adaptada à estrutura dada (uma vez que faz parte dela), nem por isso deixa de ser relativamente independente dela (BRØNDAL, 1943, p. 97) ⁴¹.

Como resultado, o sistema proposto por Brøndal permite uma abordagem dialógica simultaneamente indutiva (na qual são determinadas as direções e relações entre os termos da estrutura) e dedutiva (que constitui a base relacional da teoria a partir dos próprios termos da estrutura).

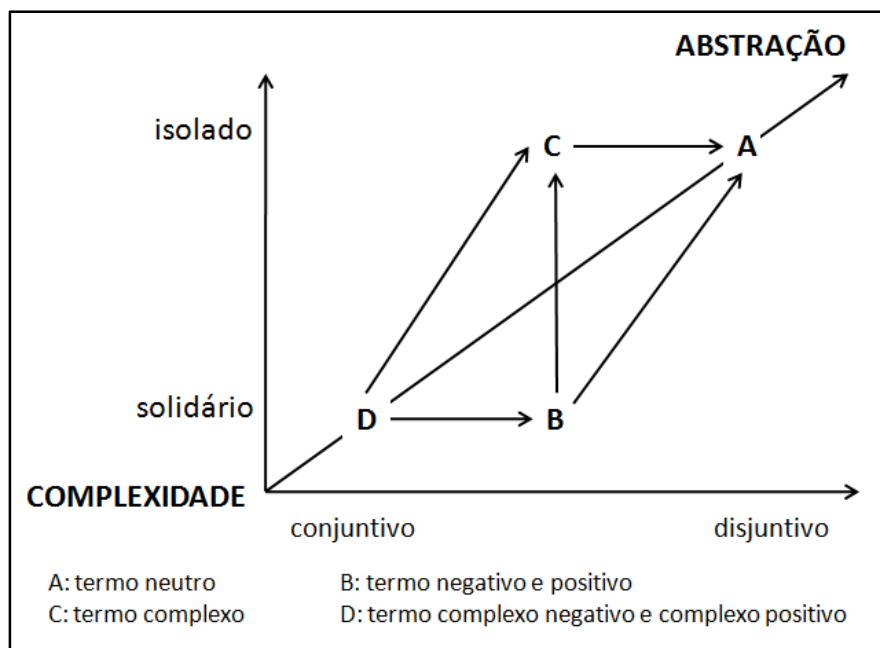
Para Brøndal, portanto, o que importa no sistema são as relações, o que existe entre um termo e outro. Na teoria do Pensamento Complexo, Morin enuncia as relações, mas, salvo engano, não explicita como operá-las. Brøndal resolve o problema definindo a noção de série, que visa a exprimir a ideia de contiguidade, um “*continuum orientado*” (ibidem). Zilberberg identificou esta mesma noção em Deleuze, quando este afirma que “o *continuum* deve ser apreendido como o fluxo de uma série contínua sensível e que do ponto de vista epistemológico, permite-se pensar que dissimetria e orientação devem entreter uma relação de pressuposição recíproca” (DELEUZE, 1968, p. 51). Nesta tese, a contiguidade das séries é operada a partir da noção de “gradação”.

Brøndal desconstrói, assim, o primado da oposição (a base da dicotomia) e simultaneamente insere o princípio recursivo na sua proposta de estrutura ao afirmar que “nos sistemas [...] duas formas polares se opõem não apenas entre si, mas ao mesmo tempo a uma forma complexa ou sintética” (BRØNDAL, 1943, p. 107) ⁴². O esquematismo do sistema elaborado por Brøndal foi representado graficamente por Zilberberg e é apresentado na Figura 6.

⁴¹ No original, em francês: “On a conçu ici la structure comme objet autonome et par conséquent comme non-dérivable des éléments dont elle n'est ni l'agrégat ni la somme ; c'est pourquoi il faut considérer l'étude des systèmes possibles et de leur forme comme étant de la plus grande importance. — Et pourtant on ne saurait considérer les éléments qui font partie d'un système comme de simples dérivés des corrélations ou oppositions structurales ; il sera en effet important de distinguer entre les propriétés purement formelles d'un système et sa matière ou substance qui, tout en étant adaptée à la structure donnée (puisque'elle y entre), n'en est pas moins relativement indépendante.” Tradução cotejada com a de Ivã Carlos Lopes, Luiz Tatit e Waldir Beividas, em Razão e Poética do Sentido (ZILBERBERG, 2006b).

⁴² No original, em francês: “Dans les systèmes où deux formes polaires s'opposent non seulement entre elles, mais en même temps à une forme complexe ou synthétique [...]” Tradução cotejada com a de Ivã Carlos Lopes, Luiz Tatit e Waldir Beividas, em Razão e Poética do Sentido (ZILBERBERG, 2006b).

Figura 6 – Esquematismo tensivo elaborado por Viggo Brøndal



Fonte: Adaptado e traduzido de (ZILBERBERG, 1987, p. 63).

Para Brøndal, a origem está na complexidade. Aquilo que nos sistemas tradicionais de classificação constitui o objetivo a ser alcançado, ou seja, a determinação do conceito dicotômico, integral e inequívoco (o termo neutro “A”), é, em Brøndal, um produto eventual das relações entre os termos advindos da complexidade.

Cabe destacar mais uma coincidência de pontos de vista a partir de Brøndal, desta vez explicitada por Morin em “A inteligência cega”, quando, em uma nota de rodapé, na qual sem citar a fonte, Morin relata:

George Lukács, o filósofo marxista, dizia na sua velhice, criticando sua própria visão dogmática: ‘O complexo deve ser concebido como o primeiro elemento existente. Daí resulta que é preciso primeiro examinar o complexo enquanto complexo e passar em seguida a seus elementos e processos elementares’ (MORIN, 2005a, p. 15).

Ao defender a complexidade como ponto de partida, Brøndal promoveu uma ruptura com o modelo saussureano, que se fundava primeiramente nas oposições entre os termos (considerava apenas as relações antagônicas). O termo complexo de Brøndal denuncia a artificialidade das oposições, uma vez que estas são produzidas a partir do amálgama, através das operações de disjunção e de isolamento dos termos. A disjunção opera na separação estrutural, através de relações simultaneamente antagônicas e concorrentes, engendrando os

termos distinguíveis “B”, ainda solidários, dependentes de suas oposições positivo/negativo. Quando os termos são isolados, acentuando o antagonismo em detrimento da concorrência, obtém-se o termo “A” neutralizado em relação a seus opostos. O termo complexo “C”, considerado por décadas o maior enigma da teoria brøndalina torna-se mais claro quando posto em diálogo com o Pensamento Complexo, pois é possível afirmar que carrega em si os princípios hologramático e recursivo, que permitem entendê-lo como uma forma de “síntese”, que só faz sentido enquanto parte da complexidade da qual emerge, e, simultaneamente, permite abordar esta complexidade a partir das suas relações com os próprios termos positivo/negativo (“D”), das quais recursivamente podem ser separados, estabelecendo o circuito complexo enunciado no “Método” de Morin e subentendido do esquema tensivo de Brøndal, ilustrado pela Figura 6.

Desta forma, além de relativizar a importância das oposições e dos termos neutros para entendimento da estrutura, Brøndal revela que o cerne do sistema tensivo são as relações entre os termos e os graus nos quais essas relações se estabelecem. Outra definição decisiva para a teoria tensiva é a afirmação de Brøndal que as relações “funcionam sempre no interior de uma mesma categoria” ⁴³ (BRØNDAL, 1943, p. 42), reafirmando a asserção de González de Gómez que os dispositivos só se definem dentro de seu “campo de operação”.

Esta afirmação conduz à definição da noção de “espaço tensivo”, aquele no qual os elementos se estabelecem a partir de suas relações; relações estas que, dialogicamente, determinam os próprios elementos (ZILBERBERG, 2006a).

Zilberberg propõe a aplicação deste modelo, nada trivial, a partir de estudos do que denomina “esquematismo tensivo” ⁴⁴, cujos fundamentos operativos podem ser resumidos da seguinte forma:

- operação por composição: o esquema tensivo opera a gradação das misturas em substituição às oposições do estruturalismo saussureano. Um conceito deixa de ser expresso em sua relação de oposição a outro conceito (por exemplo, “frio” em oposição a “quente”). Este princípio permite expressar um conceito para além das restrições da integralidade (no mesmo exemplo, “morno” como uma gradação entre “frio” e “quente”) e da inequivocidade (“morno” também pode ser simultaneamente, “um pouco frio” e/ou “um pouco quente”);

⁴³ No original, em francês: “Une opposition précise joue toujours à l'intérieur d'une même catégorie”. Tradução de Ivã Carlos Lopes, Luiz Tatit e Waldir Beividas, em *Razão e Poética do Sentido* (ZILBERBERG, 2006b).

⁴⁴ A expressão *esquematismo tensivo* foi cunhada por Zilberberg para expressar a sistematização da teoria tensiva. O autor francês prefere o seu uso em lugar de *gramática tensiva* ou *semiótica tensiva*, expressões também utilizadas por outros autores e tradutores (e pelo próprio Zilberberg em trabalhos anteriores).

- precedência da complexidade: o esquematismo tensivo considera que os termos isolados são originalmente dados juntos, e que só existem separadamente se considerados a partir das operações de composição (o que quer dizer que “frio”, “morno” e “quente” são noções derivadas de uma mesma origem complexa, e que a gradação operada pelo sensível permite juntar ou separar em termos inteligíveis separadamente) (ZILBERBERG, 2004).

Em outras palavras, partindo de uma concepção estruturalista, mas não dicotômica, o esquematismo tensivo visa a sistematizar no espaço tensivo as misturas entre duas dimensões: a do sensível e a do inteligível, que atuam sobre os conceitos. A dimensão do sensível, regente, opera pela recursividade das misturas e pela sua expressão em graus. A dimensão do inteligível, regida pela primeira, opera pela separação, pela delimitação, mas de forma não dicotômica (FONTANILLE; ZILBERBERG, 2001).

Os princípios operativos oferecidos pelo esquematismo tensivo compatibilizam formulações teóricas como o “plano de imanência” de Deleuze e o “espaço tensivo” de Zilberberg, no qual os dispositivos passam a figurar a partir das suas relações medidas em graus e das fronteiras expressas a partir das relações e não aprioristicamente, como anteviu Brøndal.

A rejeição às ideias de Brøndal é frequentemente justificada pelo ataque à sua visão estruturalista e também à sua convicção da prevalência das categorias lógicas sobre as categorias linguísticas dentro das estruturas. Neste último grupo, até Claude Zilberberg, um dos principais responsáveis pelo resgate e pela valorização da obra de Brøndal, faz coro aos críticos: “pode-se lamentar o platonismo lógico de Brøndal que o leva a situar a lógica acima das categorias linguísticas” (ZILBERBERG, 1987, p. 77. Tradução nossa)⁴⁵. A lamentação de Zilberberg pode ser justificada se considerada a aceção dicotômica da lógica clássica. No entanto, Brøndal, ao enfatizar as gradações entre as relações, parece querer fazer menção a outro tipo de lógica, incipiente em sua época, mas atualmente formulada através das lógicas não clássicas, dentre as quais figura a lógica fuzzy.

4.3 LÓGICA FUZZY

A formulação de um sistema lógico matemático para manipulação simbólica e cálculo baseado nas oposições entre os valores de “verdade” ou “falsidade” ocorreu no século XIX, a

⁴⁵ No original, em francês: “On peut encore regretter le platonisme logique de Brøndal, qui le conduit à placer la logique « au-dessus » des catégories linguistiques.”

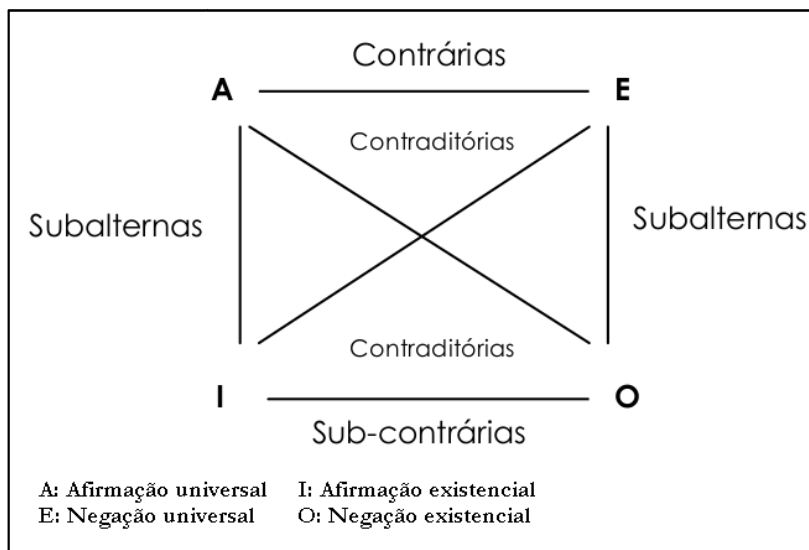
partir dos trabalhos do lógico britânico George Boole (1815-1864). Este sistema foi batizado com o nome de seu criador: a lógica booleana, que atribui 0 (zero) como valor para falso e 1 (um) como valor para verdadeiro.

A partir da lógica booleana foram desenvolvidas a lógica proposicional e a lógica de predicados. Essas três modalidades de lógica constituem o que se convencionou chamar de lógica clássica.

Na lógica clássica, as afirmações que compõem o sistema de cálculo podem receber apenas um de dois valores: verdadeiro ou falso. Além disso, as afirmações devem seguir os princípios da não-contradição (uma afirmação não pode ser simultaneamente verdadeira e falsa) e o princípio do terceiro excluído (uma afirmação ou é verdadeira ou é falsa, não havendo outra possibilidade).

O quadrado das proposições, cuja concepção é creditada a Aristóteles, articula os princípios da lógica clássica em uma representação esquemática, conforme ilustra a Figura 7.

Figura 7 – Quadrado das proposições, síntese das operações na lógica clássica



Fonte: Adaptado a partir do verbete “*The Traditional Square of Opposition*”, na Enciclopédia de Filosofia Stanford (PARSONS, 2015).

Deste sistema de cálculo, é possível obter os seguintes resultados:

- A e O; e I e E são contraditórias. Não podem ser ambas verdadeiras; e não podem ser ambas falsas.
- A e E são contrárias. Não podem ser ambas verdadeiras; mas podem ser ambas falsas

- I e O são subcontrárias. Não podem ser ambas falsas; mas podem ser ambas verdadeiras.
- I é subalterna de A e O é subalterna de E. Se A é verdadeira, então I é verdadeira. Se E é verdadeira, então O é verdadeira.

Com o chamado “renascimento da lógica” a partir do século XIX, a suficiência do modelo aristotélico começou a ser questionada, e, já no século XX, os princípios do que viria a se tornar a lógica fuzzy foram desenvolvidos pelo lógico polonês Jan Łukasiewicz (1878-1956), que estendeu a lógica booleana, adicionando um terceiro valor possível entre o 0 e o 1, o $\frac{1}{2}$, que indicaria “nem verdadeiro, nem falso” (ibidem). Posteriormente, Łukasiewicz expandiu os valores intermediários para o conjunto infinito de números reais entre 0 e 1 que passaram a representar graus de verdade, como, por exemplo, 0,9 para quase verdadeiro e 0,1 para quase falso.

A formulação e a denominação “lógica fuzzy”, entretanto, surgiu em 1965, quando Lofti Asker Zadeh formalizou os conjuntos fuzzy ⁴⁶ e demonstrou matematicamente as suas regras de operação:

Um conjunto fuzzy é uma classe de objetos com graus de pertencimento. Tal conjunto é caracterizado por uma função de pertencimento que atribui a cada objeto um grau que varia entre zero e um. As noções de inclusão, união, intersecção, complemento, relação, convexidade, etc., são estendidas a esses conjuntos. ⁴⁷ (ZADEH, 1965, p. 338. Tradução nossa).

Zadeh reconhece o papel de sua invenção em relação ao desafio da complexidade, e defende que o pensamento fuzzy, que sofria considerável rejeição por parte dos matemáticos, “não deveria ser tão atacado, pois ele tornava possível a análise de problemas complexos” (ZADEH, 1971).

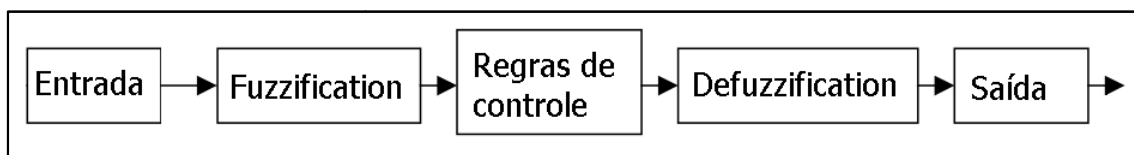
Esta defesa, feita por Zadeh há várias décadas, é pertinente ainda hoje, pois parte significativa das críticas em relação à lógica fuzzy decorre do fato de se considerá-la indevidamente uma “teoria do conhecimento”, quando é “apenas” uma teoria matemática sem pretensão interpretativa ou semântica. Ou seja, a lógica fuzzy não visa a fornecer um modelo para a determinação de conceitos, é uma ferramenta de cálculo, que opera a partir de representações lógicas utilizando princípios matemáticos (BĚLOHLÁVEK *et al.*, 2009).

⁴⁶ No original, em inglês, “fuzzy sets”.

⁴⁷ No original, em inglês: “A fuzzy set is a class of objects with a continuum of grades of membership. Such a set is characterized by a membership function which assigns to each object a grade of membership ranging between zero and one. The notions of inclusion, union, intersection, complement, relation, convexity, etc., are extended to such sets”.

A primeira aplicação da teoria fuzzy para além dos experimentos académicos foi concluída em 1974, tendo sido realizada por Ebrahim H. Mamdani (1942-2010), que desenvolveu um controle de uma máquina a vapor, baseado em lógica fuzzy (MAMDANI; ASSILIAN, 1975). A aplicação da lógica fuzzy se dá a partir de um processo previamente estabelecido, que utiliza descrições linguísticas em linguagem natural em lugar das equações diferenciais tradicionalmente utilizadas para construir modelos matemáticos de fenômenos físicos. Esta característica permite o aproveitamento da noção heurística contida nestas descrições (chamadas operadores fuzzy), sendo, portanto, uma simplificação deliberada da própria teoria fuzzy, e que torna a representação lógica mais intuitiva e o processamento computacionalmente realizável (KLIR; ST. CLAIR; YUAN, 1997). O processo fuzzy adotado nesta tese é o mesmo proposto por Mamdani, e que recebe seu nome, e está representado na Figura 8:

Figura 8 – Etapas de um processamento fuzzy, segundo o processo Mamdani



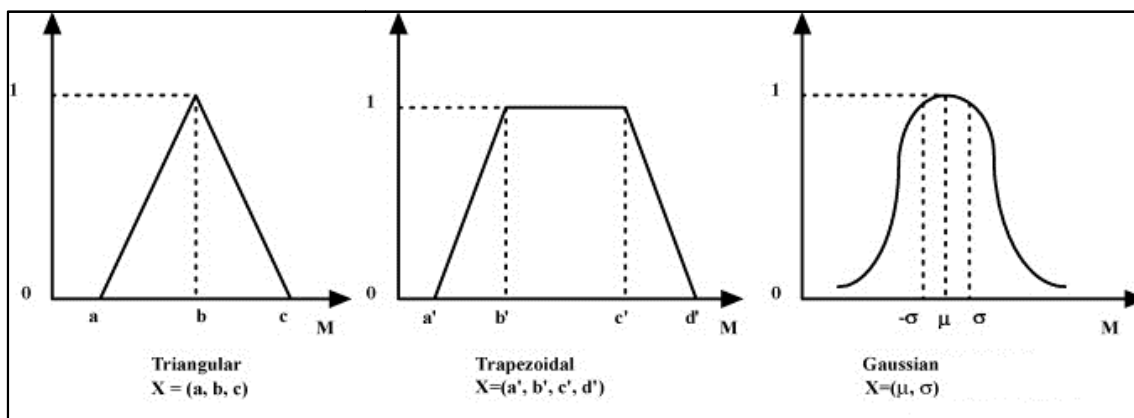
Fonte: Elaborado pelo autor.

Para ilustrar cada etapa do processamento fuzzy, será utilizado um exemplo aparentemente trivial: a abordagem ao conceito “morno” em uma hipotética categorização de temperaturas de líquidos. Em um sistema de classificação, seria necessário definir as fronteiras fixas do conceito “morno”, por exemplo, determinando como “morno” os valores entre 25°C e 45°C. Isso significa, seguindo os princípios da integralidade e da inequivocidade que 24°C não seria absolutamente morno, tanto quanto não seriam 0°C ou 90°C.

- As entradas de um sistema fuzzy, são expressas em variáveis linguísticas, no caso, o conceito “morno”. A entrada é constituída também por um valor ou conjunto de valores que se deseja calcular. No exemplo, uma temperatura hipotética de 24°C.

- A etapa de *fuzzification*⁴⁸ transforma as entradas de termos linguísticos e variáveis escalares em variáveis fuzzy. Há diferentes possibilidades de representação analítica da função de *fuzzification*, conforme ilustra a Figura 9:

Figura 9 – Tipos de função de função de *fuzzification*: triangular; trapezoidal e Curva de Gauss



Fonte: Adaptado e traduzido de (KLIR; ST. CLAIR; YUAN, 1997, p. 80–86).

A diferença fundamental entre os tipos de funções de *fuzzification* diz respeito ao cálculo do grau de pertencimento de um dado conceito fuzzy. Na função triangular, o pertencimento total só ocorre no vértice superior do triângulo (valor “b”). Nas demais áreas do triângulo, o grau de pertencimento é parcial (inicialmente crescente a partir de “a”, e, depois do vértice “b” é decrescente até o valor “c”). Na função trapezoidal, há uma determinada gama de valores para os quais o pertencimento é total (na figura, os valores entre “b’” e “c’”). A função curva de Gauss permite uma gama maior e mais gradual de níveis de pertinência, com o pertencimento total restrito a uma determinada faixa – representada na Figura 9 pelos valores arbitrários Θ (*theta*) e υ (*upsilon*) – sendo a função mais adequada para representar os graus de pertencimento a “morno”, no exemplo considerado.

- As regras de controle são os construtos lógicos que operam sobre as variáveis fuzzy, através de operações das teorias dos conjuntos definidas por Zadeh (inclusão, união, intersecção, complemento, relação, convexidade) e geram os

⁴⁸ Os termos decorrentes das ações fuzzy, são frequentemente traduzidos como “fuzzificação” e “desfuzzificação”. Para evitar a adoção de tais neologismos, e mantendo a coerência em relação à adoção do termo original em inglês fuzzy, estes termos serão mantidos sem tradução, porém, destacados em itálico.

valores resultantes do cálculo fuzzy e que servem como entrada para a etapa de *defuzzification*. No caso do exemplo, por haver apenas uma variável, as regras se limitam a, de acordo com o tipo de função escolhida, devolver o valor adequado de pertencimento do valor “24°C” em relação ao termo “morno”.

- A etapa de *defuzzification* é responsável pelo processo de transformação dos valores fuzzy resultantes da aplicação das regras de controle novamente em variáveis linguísticas, calculando o respectivo grau de pertencimento. Existem diversos métodos de *defuzzification*. O mais utilizado nos sistemas fuzzy é o método de centro de gravidade, que calcula a área da curva da variável linguística de saída produzida pela aplicação das regras e identifica o índice correspondente que divide esta área pela metade. No exemplo adotado, o retorno do conceito “morno” para a entrada 24°C teria um grau de pertencimento, por exemplo, 0,7. Supondo a existência de outro conceito expresso pelo termo linguístico “quente”, possivelmente, seu grau de pertencimento seria 0 (zero).
- A saída apresenta as variáveis linguísticas resultantes da *defuzzification*, com seu respectivo grau de pertencimento expresso em valores percentuais (de 0 a 100). Ainda em relação ao exemplo da temperatura, a entrada 24°C teria uma pertinência de 70% em relação ao termo linguístico “morno” e de 0% ao termo linguístico “quente”.

A aplicação de regras fuzzy não resolve verdadeiramente o problema da dicotomia, uma vez que continua a operar sobre as oposições de valores. A diferença em relação à lógica clássica é que ela permite a multiplicação de “subdicotomias”, ou seja, de estados intermediários que podem ser utilizados para representar a noção de gradação. É possível operar este tipo de representação em sistemas informatizados graças à evolução das técnicas e ferramentas computacionais que vem ocorrendo nas últimas décadas, e que são discutidas na seção seguinte.

4.4 COMPUTAÇÃO, WEB E DOCUMENTO

A computação vem se desenvolvendo a uma velocidade vertiginosa, a ponto de tecnologias serem criadas e suplantadas em questão de meses ou anos. Este preâmbulo é necessário, especialmente porque algumas das discussões apresentadas estão relacionadas a uma das áreas mais novas e imprevisíveis da computação: a web. Contudo, antes de chegar à web, é conveniente situar o presente trabalho em relação às bases suportam estas tecnologias: a computação em geral e, novamente, os documentos.

Os computadores são máquinas digitais, compostas por milhões (em alguns casos bilhões) de blocos lógicos, que, por sua vez, são constituídos por circuitos lógicos. Um circuito lógico é a menor unidade de processamento de um computador e opera internamente com dois níveis de tensão (pulsos elétricos) que são convertidos para níveis lógicos correspondentes aos valores binários 0 e 1. Em suma, tudo o que um computador digital é capaz de processar, em última instância é um valor booleano, uma perfeita dicotomia, chamada “bit”, do inglês *Binary Digit* (TANENBAUM, 2006).

Apesar de serem máquinas binárias, o alto poder computacional reside na quantidade de bits que um computador pode processar simultaneamente e na velocidade com que é capaz de realizar estes processamentos. Atualmente, o que limita esta velocidade é uma restrição física, pois as operações de manipular pulsos elétricos geram calor que não consegue ser dissipado na mesma velocidade em que é gerado, mesmo quando utilizados materiais semicondutores como o silício. Por este motivo, os fabricantes de processadores vem sendo obrigados a limitar a velocidade máxima dos processadores (chamado *clock*). Como não é possível para a Computação revogar as leis da Física, a principal aposta da indústria e da ciência para superar esta limitação vem sendo a computação quântica, ainda em fase experimental, mas que se baseia na substituição do bit binário pelo bit quântico. O bit quântico, por sua vez, objetiva justamente superar a dicotomia binária, pois permitirá os valores 0, 1 ou uma sobreposição quântica destes valores ⁴⁹.

Como insiste Edgar Morin em seu discurso sobre a necessidade de “reformatar o pensamento”, a ciência de base cartesiana, que desde o século XVII apostou na simplificação

⁴⁹ A sobreposição quântica, por sua vez, é princípio fundamental da Mecânica Quântica que afirma que um sistema físico existe parcialmente em todos os estados teoricamente possíveis simultaneamente antes de ser medido. Para uma introdução à Mecânica Quântica, ver: SAKURAI, Jun John; NAPOLITANO, Jim. *Modern quantum mechanics*. New Jersey: Pearson, 2013.

e na dicotomia, vem dando provas de esgotamento, e novas alternativas precisam ser encontradas (MORIN, 2010).

Apesar de o cenário técnico hoje disponível (ainda) ser o de computadores binários, a computação vem sofrendo uma profunda modificação nas últimas décadas, superando a velha acepção de “processamento de dados” para oferecer ferramentas capazes de executar tarefas complicadas a partir de sistemas computacionais cada vez mais sofisticados. Um destes sistemas é a *world wide web*, *www*, ou simplesmente *web*, um sistema de documentos em *hipermídia* que são interligados e executados na Internet, a rede mundial de computadores.

A *web* é um desenvolvimento consideravelmente recente, que colocou em prática ideias que até aquele momento não pareciam realizáveis. Sua criação é tributada ao físico inglês Tim Berners-Lee, que lançou em 1989 um projeto intitulado “Proposta de Gestão de Informação”⁵⁰ e que consistia em um utilizar documentos digitais de hipertexto para o compartilhamento de informações entre os pesquisadores do Centro Europeu para a Pesquisa Nuclear (CERN)⁵¹. O projeto prosperou e, notadamente a partir de 1994, se popularizou e alcançou escala mundial. Naquele mesmo ano, Tim Berners-Lee conduziu a criação do W3C (*World Wide Web Consortium*), um consórcio internacional que agrega empresas, órgãos governamentais e organizações independentes com a finalidade de estabelecer os padrões para a criação e a interpretação de conteúdos para a *web*, garantindo a compatibilidade entre as tecnologias, linguagens e aparelhos físicos (celulares, computadores etc.).

Estudos históricos demonstram a anterioridade da noção de hipertexto ao surgimento da *web* (CHARTIER, 1988), (BURKE; BRIGGS, 2004). Desde a Bíblia, passando pelo MEMEX de Vannevar Bush⁵² em 1945 ou o conceito de “*Lexia*” de Roland Barthes⁵³ em 1970, é seguro dizer que a noção foi apropriada pela *web* e colocada em prática em escala global. Sucintamente, o hipertexto consiste em um texto ao qual se agregam outros conjuntos de dados (que podem ser outros textos, imagens ou sons) e cujo acesso se dá através de indicações de referências explícitas, que, no caso da *web* são denominadas *hyperlinks*.

Apontada como a realização do “sonho de Paul Otlet” (PINHEIRO; PEREIRA, 2000), a *web* se configura como uma rede de “interligação do conhecimento em escala global” cuja

⁵⁰ Do original, em inglês, “*Information Management: a proposal*”. Disponível em: <<https://www.w3.org/History/1989/proposal.html>>. Acesso em: abr 2016.

⁵¹ Do original, em francês, *Centre Européen pour la Recherche Nucléaire*.

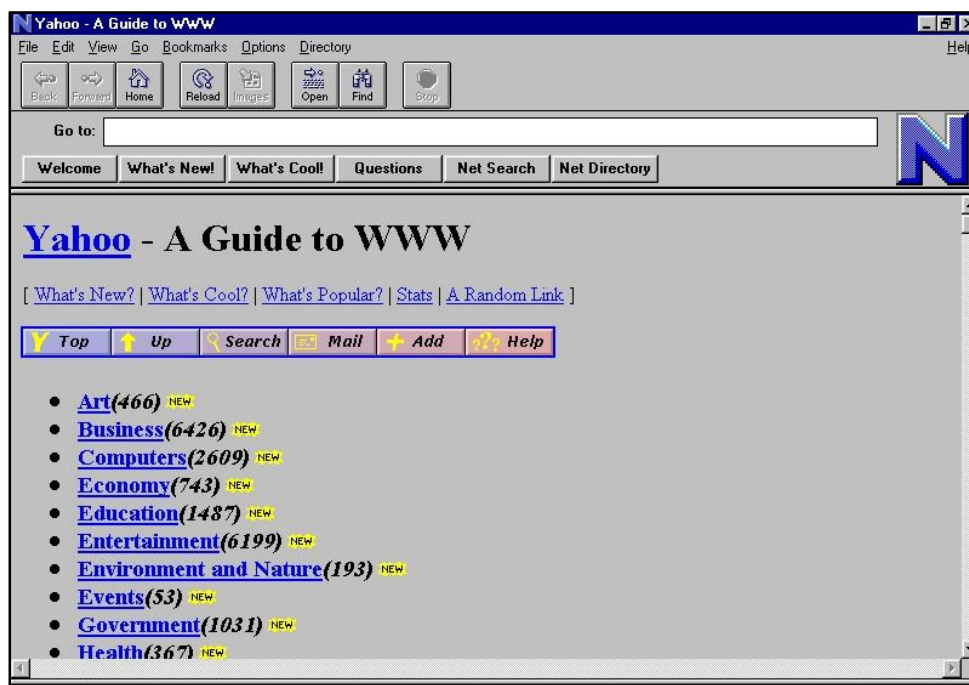
⁵² Vannevar Bush (1890-1974) foi um engenheiro, inventor e político estadunidense, conhecido pelo seu apoio político na criação da bomba atômica e pela ideia do *memex*, considerado um precursor da *web*.

⁵³ Roland Barthes (1915-1980) foi um escritor, sociólogo, crítico literário, linguista, semioticista e filósofo francês.

fruição, no entanto, é restrita àqueles que possuem a “competência digital adequada” (BARRETO, 2008). A totalidade da rede é composta pela teia de documentos ligados, retomando-se, portanto, a tradição da documentação do século XX que considera o documento como uma “matriz na qual a informação e o conteúdo são formatados no plano comunicacional”, sendo ele, ao mesmo tempo, “o suporte que permite a sua circulação” (COUZINET, 2008, p. 57).

Inicialmente a web continha apenas documentos de hipertexto de conteúdo estático (textos, hiperlinks fixos e imagens). Todo o conteúdo das páginas era somente para leitura, não havendo interação do utilizador na produção dos conteúdos. Somente o programador (chamado *webmaster*) poderia realizar alterações no conteúdo ou promover atualizações das páginas (BERNERS-LEE; FISCHETTI, 1999). A Figura 10 mostra uma página típica deste período, a página principal do site “Yahoo!” em 1994, cuja intenção era catalogar todos os sites existentes, agrupando-os em categorias.

Figura 10 – Página web da “Yahoo!” em 1994.



Fonte: W3C – The Open Web Platform History

No início dos anos 2000, as tecnologias da web passaram a permitir a interação dos utilizadores, configurando o que se convencionou denominar “web social” ou “web colaborativa”. Os conteúdos deixaram de ser estáticos e os próprios utilizadores passaram a

dispor das ferramentas disponíveis na plataforma web para produzir conteúdos (como os blogs, as mídias sociais, a Wikipedia etc.). Esta evolução levou ao crescimento exponencial do volume de dados publicado na web. Contudo, os conteúdos tornados disponíveis na web continuam a ser criados para serem interpretados exclusivamente pelos seres humanos. Um exemplo de sistema da web colaborativa é o canal de vídeos YouTube ⁵⁴, ilustrado através da Figura 11, com o conteúdo de um vídeo da canção “A Mulata é a Tal”.

Figura 11 – Página do YouTube, vídeo para a canção “A Mulata é a Tal”



Fonte: Adaptado do site YouTube (suprimidas as informações específicas do canal).

⁵⁴ Site para compartilhamento de vídeos em formato digital. Disponível em: <<http://www.youtube.com>>. Acesso em: mai 2016.

A afirmação que as páginas da web são construídas para serem interpretadas por seres humanos pode ser confirmada a partir de uma breve análise da Figura 11. Um ser humano é capaz de compreender e identificar os dados e as relações nela contidos, por exemplo:

- A Mulata é a Tal é o título de uma música.
- Os compositores desta música cujo título é A Mulata é a Tal são João de Barro e Antonio Almeida. Isto é possível descobrir porque no texto consta `marcha de, e o "de" quer dizer autoria`. Um ser humano sabe que o autor de uma música é uma pessoa. Esta pessoa é chamada de João de Barro. Além disso, o "e" indica que há mais de um autor, portanto que, além de João de Barro, Antonio Almeida, que também é presumivelmente uma pessoa, é compositor da mesma música, o que, constitui uma parceria entre João de Barro e Antonio Almeida.
- Ao ver o vídeo, um ser humano é capaz de identificar que no vídeo aparece o rosto de uma mulher.
- Ao ler o texto explicativo, um ser humano é capaz de descobrir que esta mulher se chama Maria D'Aparecida e relacionar que o rosto que aparece no vídeo é o de Maria D'Aparecida.
- Finalmente, um ser humano é capaz de relacionar que a música, cujo título é A Mulata é a Tal, cujos compositores são João de Barro e Antonio Almeida foi dedicada à Maria D'Aparecida.

Um computador não consegue realizar este tipo de associações, pois os dados que constituem os documentos na web são sequências de textos dispostas sintaticamente, sem explicitar a semântica dos seus relacionamentos. As consequências desta limitação são determinantes para o uso que se pode fazer da web. Por exemplo, a mesma ferramenta (YouTube) utilizada para encontrar a música pelo título (único campo indexado), não consegue recuperar músicas inspiradas em mulheres ou parcerias entre João de Barro e Antonio Almeida.

Em suma, atualmente os conteúdos são organizados na web de forma que apenas o ser humano é capaz de interpretá-los e, conseqüentemente, conduzir a tarefa de recuperação. Com a web, a tarefa de recuperação, preterida durante décadas, vem se tornando central para os sistemas de organização do conhecimento.

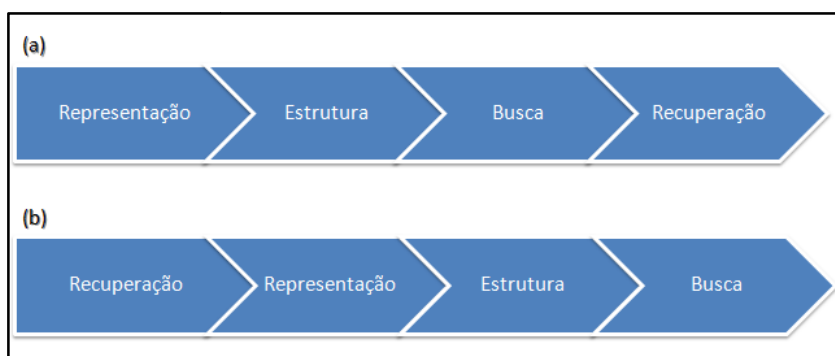
4.5 ORGANIZAÇÃO DO CONHECIMENTO DIRIGIDA PELA RECUPERAÇÃO

A proposta de uma Organização do Conhecimento dirigida pela recuperação não é uma novidade advinda da web e das novas tecnologias. Jesse Shera (1903-1982) defendia desde a década de 1950 que “a recuperação deveria ser o foco” de qualquer teoria da organização do conhecimento e, portanto, “o objetivo em direção ao qual todos os nossos esforços devem ser encaminhados” ⁵⁵ (SHERA, 1960, p. 136). Esta visão é corroborada pelo argumento que a recuperação é o objetivo final de um sistema de organização do conhecimento e o único do qual os indivíduos que utilizam os sistemas efetivamente participam (ibidem).

No entanto, na visão tradicional, usualmente é a representação que constitui a primeira etapa do processo de criação dos sistemas de organização do conhecimento. A representação dos recursos (o título da canção no exemplo do vídeo no Youtube) determina e restringe as estruturas de organização que, por sua vez, determinam as possibilidades de busca, que, transitivamente, determinam o conjunto de recursos que serão recuperados pelo sistema.

Shera enfatizava o papel crítico da recuperação nos sistemas de organização do conhecimento e postulava que fosse ela, a recuperação, a etapa determinante das estruturas de representação e de busca adotadas (SHERA, 1956). Na Figura 12, a parte “(a)” ilustra as etapas da metodologia tradicional de sistemas de organização do conhecimento, enquanto a parte “(b)” ilustra o deslocamento proposto por Shera.

Figura 12 – Etapas de um sistema de organização do conhecimento dirigido pela representação (a) e pela recuperação (b)



Fonte: Elaborado pelo autor.

⁵⁵ No original, em inglês: “Therefore, [...] retrieval, must be the focal point of a library theory and the end toward which all our efforts are directed.”.

A proposição de Shera vem se materializando como metodologia adotada na web e caracteriza o já mencionado deslocamento paradigmático da Organização do Conhecimento hierárquica (as árvores) para uma concepção relacional (as redes e rizomas). Apesar de enfrentar resistências dentro alguns seguimentos da Organização do Conhecimento (como qualquer mudança paradigmática enfrenta), as soluções baseadas no princípio da recuperação, e cujo exemplo extremo é o da ferramenta Google ⁵⁶ – uma caixa de texto livre, a partir da qual é possível operar uma infinidade de requisições –, convidam a Organização do Conhecimento a revisitar, além de Shera, outro de seus grandes teóricos, o já aludido Brian Vickery, que com sua admirável visão holística advertia em 1986 que os sistemas de organização do conhecimento deviam ser pensados para levar em conta primeiramente a recuperação, mas não apenas a recuperação manual, devendo considerar também a possibilidade de raciocínio automático – aquele realizado pelo próprio computador utilizando representações e estruturas internas de processamento (VICKERY, 1986). Segundo Vickery, a possibilidade de raciocínio computacional redefiniria as estratégias de buscas “direta” e “por navegação” ⁵⁷ para técnicas automatizadas (VICKERY, 2004), o que atualmente vem sendo precisamente o investimento em uma nova abordagem para a web: a web semântica.

4.6 WEB SEMÂNTICA

O projeto de construção da web semântica vem sendo liderado também por Tim Berners-Lee. Em 1999, o cientista inglês publicou um livro no qual fez uma avaliação da web até aquele momento, diagnosticando que a web era incompleta. Esta incompletude se devia ao fato de os recursos ⁵⁸ nela existentes não serem exatos nem processáveis por computadores. No mesmo texto, Berners-Lee apontava possíveis desdobramentos para o futuro próximo. Uma das ideias lançadas foi que os conteúdos na web deveriam ser integrados, de forma que os computadores pudessem “compreender” a necessidade de recuperação informada pelos utilizadores e serem capazes de retornar dados a partir do cruzamento de recursos, de forma

⁵⁶ Site de pesquisas Google, da empresa de mesmo nome. Disponível em: <<http://www.google.com>>. Acesso em: mai 2016.

⁵⁷ Tradução nossa para “*seeking and browsing*”.

⁵⁸ Esclarecimento terminológico: na computação o termo “recurso”, do inglês “*resource*”, é largamente utilizado para se referir genericamente aos objetos (coisas, conjuntos de dados, páginas etc.). Nesta seção, será preservado o uso do termo genérico “recurso”, porém, cabe adiantar que, no momento da sua aplicação, um “recurso” corresponderá a um “dispositivo”.

estruturalmente análoga ao que os seres humanos fazem (BERNERS-LEE; FISCHETTI, 1999).

No entanto, foi em 2001, em um artigo publicado na revista *Scientific American*, que o inventor da web cunhou a expressão “web semântica” e a definiu não como uma nova web, mas uma extensão da atual. Nas palavras de Berners-Lee, “uma nova forma de conteúdos na web compreensíveis pelos computadores [que] desencadeará uma revolução de novas possibilidades”⁵⁹ (BERNERS-LEE; HENDLER; LASSILA, 2001, p. 28. Tradução nossa).

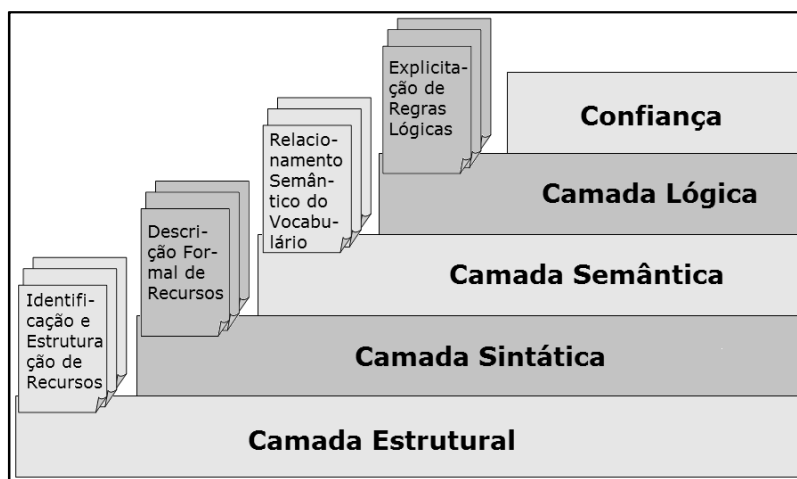
A partir desta constatação, foi criado um projeto de grande escala, sempre conduzido pelo consórcio W3C, que tem por objetivo a proposta de padrões para se atingir a web semântica, baseado nos seguintes princípios:

- (a) Os recursos disponíveis na web devem ter significado exato;
 - (b) Os recursos disponíveis na web devem ser processáveis por computadores;
 - (c) Os computadores devem poder integrar de forma automática os recursos da web
- (SEMANTIC WEB, 2015).

Por não se tratar de uma nova web, a web semântica vem sendo desenvolvida conjuntamente à web “atual”, seguindo a ideia de camadas, que são sobrepostas aos recursos já existentes, no intuito de lhes atribuir semântica e garantir o cumprimento dos três princípios supramencionados. Esta visão em camadas é um dos grandes trunfos da web semântica em relação a outras iniciativas globais, pois, com o volume atual (e sempre crescente) de dados publicados na web, qualquer projeto que pressupusesse uma interrupção ou a necessidade de processamento prévio completo (como a catalogação de tudo o que já existe) seria irrealizável. Esta arquitetura em camadas é ilustrada pela Figura 13.

⁵⁹ Do original, em inglês, “A new form of web content that is meaningful to computers will unleash a revolution of new possibilities”.

Figura 13 – Arquitetura da web semântica



Fonte: Adaptado e traduzido do site W3C.

A seguir é apresentada uma breve síntese a respeito de cada uma das camadas da arquitetura da web semântica, sendo que as camadas semântica e lógica, por serem as mais relevantes para este trabalho, serão pormenorizadas logo a seguir.

- **Camada estrutural:** diz respeito à identificação dos recursos na web. Todo recurso deve ser persistente e univocamente identificado através de um URI - Identificador Uniforme de Recursos ⁶⁰. Usualmente o URI é composto pelo endereço do site no qual está hospedado (chamado de domínio), seguido de um código único. No exemplo apresentado anteriormente, o URI da canção “A Mulata é a Tal” no Youtube é <https://www.youtube.com/video/iTIrCFdgPpE>. Cabe ressaltar que o URI não necessariamente equivale a um URL – Localizador Uniforme de Recursos ⁶¹ – sendo este último um tipo específico de URI que corresponde a um endereço na web que pode ser acessado por um utilizador (em resumo, todo URL é um URI, mas o contrário não é necessariamente verdade). Esta primeira restrição responde ao princípio “(a)” descrito anteriormente, uma vez que cada recurso disponível na web será referenciado univocamente, sendo possível ao computador identificá-lo de forma exata. Mais uma vez recorrendo ao exemplo, o computador será

⁶⁰ Sigla URI, do original, em inglês, *Uniform Resource Identifier*.

⁶¹ Sigla URL, do original, em inglês, *Uniform Resource Locator*.

capaz de entender que <https://www.youtube.com/video/iTirCFdgPpE> corresponde àquele vídeo específico, com a gravação de “A Mulata É A Tal”.

- **Camada sintática:** diz respeito à descrição dos recursos, usualmente realizada através de metadados, expressos na forma “atributo / valor”, conforme exemplificado no Quadro 1.

Quadro 1 – Exemplo de metadados para a canção “A Mulata é a Tal”

Atributo	Valor
Título	A Mulata é a Tal
Intérprete	Ruy Rey
Ano de gravação	1947

Fonte: Elaborado pelo autor.

O uso de metadados na web enfrenta três problemas principais: o primeiro diz respeito à falta de padronização no preenchimento, uma vez que diferentes culturas ou mesmo diferentes indivíduos ou organizações podem fazer usos distintos do preenchimento de nomes próprios ou de datas, por exemplo. O segundo problema refere-se à falta de cobertura, pois um conjunto, por mais extenso que seja, usualmente não cobre todas as especificidades do domínio. Nestes casos, ou o padrão é estendido (deixando de ser “padronizado”), ou os dados precisarão ser adaptados em campos que não eram originalmente previstos para aquele conteúdo. O terceiro problema refere-se à falta de semântica, ou seja, de explicitação do que os campos e seus valores significam. Da mesma forma que um documento web, um padrão de metadados é legível por humanos, mas os computadores não são capazes de associar as relações implícitas nos pares “atributo / valor”.

- **Camada semântica:** é a responsável por atender ao princípio “(b)”, ou seja, fazer com que os recursos na web sejam processáveis por computadores. Nesta camada são definidos os relacionamentos entre os recursos, explicitando a sua semântica. Esta camada será pormenorizada a seguir, incluindo a apresentação de suas linguagens, o RDF e a OWL.
- **Camada lógica:** é a responsável pelo item “(c)”, terceiro e último princípio da web semântica, que tem por objetivo permitir que o computador seja capaz de

“raciocinar” ⁶² a partir dos recursos e de seus relacionamentos. Este “raciocínio” corresponde precisamente ao que Brian Vickery antevia como substituição das estratégias de buscas para técnicas automatizadas. Neste trabalho, esta camada é alcançada com a utilização combinada entre a linguagem OWL e as propriedades fuzzy, conforme detalhado a seguir.

- **Camada de confiança:** tem por objetivo definir mecanismos de controle e garantia da confiabilidade das buscas automáticas realizadas na camada lógica. Sua principal intenção é evitar os chamados “sofismas computacionais” e as inferências ambíguas. Como só faz sentido após as camadas semânticas e lógicas estarem devidamente consolidadas, não será formalmente abordada no âmbito deste trabalho.

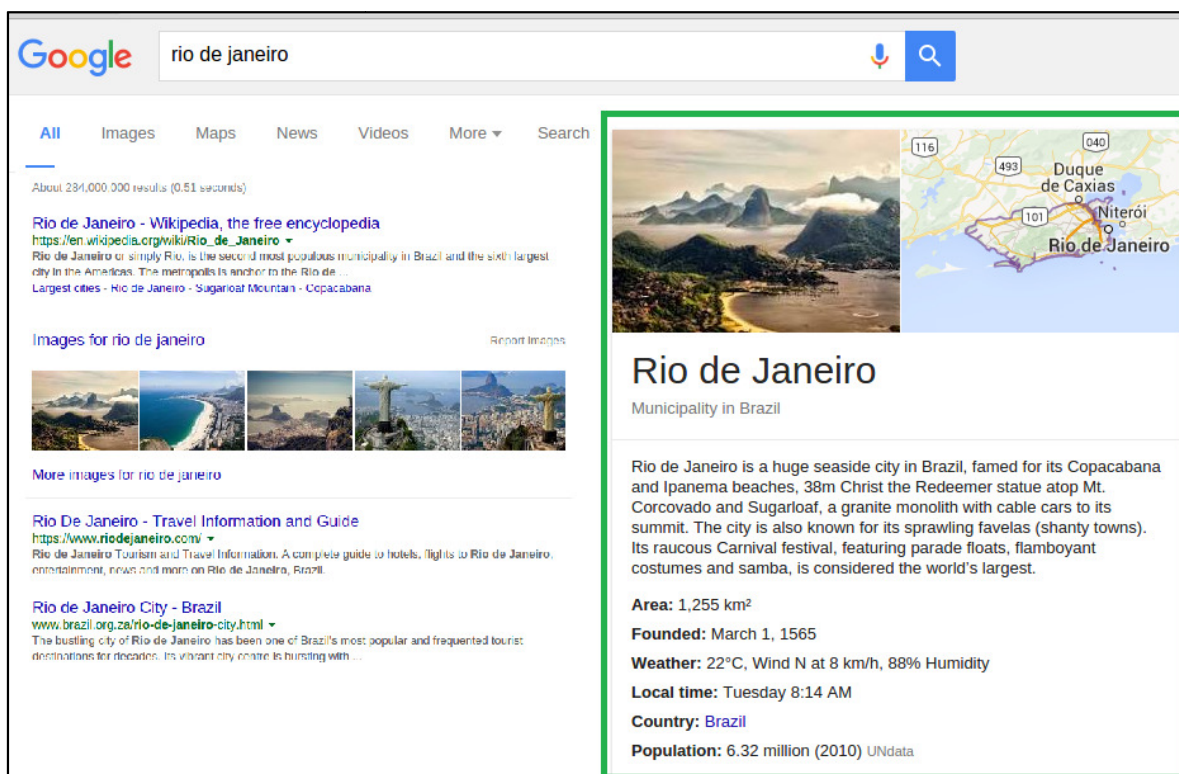
Linguagem OWL – Semântica e lógica

O projeto da web semântica consiste em transformar os recursos existentes na web em artefatos computacionalmente processáveis, permitindo o “raciocínio” automático e a integração automática destes conteúdos pelos computadores a partir de regras lógicas. A sua consecução ainda é parcial. Sabidamente trata-se de um projeto de longa duração e de permanente evolução, porém resultados parciais tem se mostrado promissores (AGHAEI; NEMATBAKHSH; FARSANI, 2012), (CHUNG, 2015).

Um exemplo possivelmente familiar de sua utilização é quando se procura o nome de uma pessoa, de um lugar ou de uma instituição no sistema de pesquisas Google. No exemplo apresentado na Figura 14, é ilustrada a pesquisa pela expressão “Rio de Janeiro”.

⁶² Do termo corrente em inglês, *reasoning*.

Figura 14 – Resultados obtidos por tecnologias da web semântica em pesquisa no Google



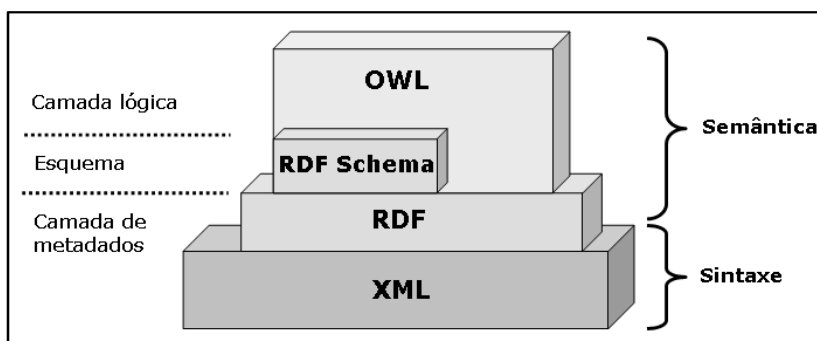
Fonte: Adaptado do site Google, enfatizando os dados retornados com recursos da websemântica (retângulo ao lado direito da imagem).

Além dos resultados usualmente retornados por uma ferramenta de pesquisa, que consiste na lista de sites apresentada no lado esquerdo da Figura 14, a ferramenta é capaz de “entender” que a expressão Rio de Janeiro refere-se a uma cidade e, na coluna do lado direito, apresenta uma foto e o mapa desta cidade. Além disso, apresenta informações extraídas da Wikipedia⁶³ como uma descrição, a data de fundação da cidade e dados em tempo real como a hora local, a temperatura, a velocidade do vento e o percentual de umidade, obtidos de um serviço meteorológico online (*Weather Underground*) (KEITH, 2003).

Como discutido anteriormente, estas associações, embora sejam imediatas para os seres humanos, precisam ser tornadas explícitas para o computador, através de descrições semânticas. Estas descrições são realizadas em três níveis de profundidade, conforme ilustra a Figura 15.

⁶³ Para ser preciso, os dados da Wikipédia para a web semântica fazem parte do projeto DBpedia <<http://dbpedia.org>>, que aplica uma camada semântica sobre cópias de alguns conteúdos da Wikipédia Acesso em: mai 2016.

Figura 15 – Níveis de descrição na web semântica



Fonte: Adaptado e traduzido do site W3C.

O primeiro nível é alcançado pela utilização da linguagem XML ⁶⁴ – Linguagem de Marcação Extensível. A XML é uma linguagem estruturalmente simples e flexível, restrita a descrições sintáticas. Através dela é possível representar conjuntos heterogêneos de dados, organizando-os através da representação através de pares atributo / valor, de forma hierárquica, como ilustra o trecho de Código fonte 1.

Código fonte 1 – Dados sobre a cidade Rio de Janeiro, expressa em XML

```
<cidade>
  <nome>Rio de Janeiro</nome>
  <fundacao>1/3/1565</fundacao>
  <fundador>Estácio de Sá</fundador>
</cidade>
```

Fonte: Elaborado pelo autor.

A XML constitui a base para as demais linguagens da web semântica, sendo uma delas o RDF – *Framework* para Descrição de Recursos ⁶⁵. Este *framework* ⁶⁶ apresenta uma importante diferença: enquanto a XML define a estrutura dos dados na forma atributo / valor, o RDF os declara na forma de triplas: sujeito / predicado / objeto.

⁶⁴ Do inglês *eXtensible Markup Language*.

⁶⁵ Do inglês, *Resource Description Framework*

⁶⁶ *Framework* é uma palavra da língua inglesa que em tradução literal quer dizer “armação”, “arcabouço”. No âmbito da Computação, refere-se a uma abstração que tem o intuito de prover o aproveitamento de diversas funcionalidades genéricas para reutilização em uma nova aplicação. No entanto, as palavras “arcabouço” ou “armação” não são correntes na literatura científica em língua portuguesa para este contexto, preferindo-se, portanto, utilizar o termo em inglês *framework*.

Os sujeitos são os elementos que se deseja descrever, como, por exemplo, *Cidade*, *Fundador*, *Clima*. Os objetos são os elementos que contêm os valores das descrições, como *Nome* e *Data de fundação*. Os predicados são as relações entre sujeitos e objetos, como *possui nome*, *fundada por*, *pertence a* etc., exemplificadas no Código fonte 2.

Código fonte 2 – Dados de cidade e fundador, expressos em RDF

```
<rdfs:Class rdf:ID="cidade"/>
<rdfs:Class rdf:ID="fundador"/>
<rdfs:ObjectProperty rdf:ID="fundadaPor">
  <rdfs:domain rdf:resource="#cidade"/>
  <rdfs:range rdf:resource="#fundador"/>
</rdfs:ObjectProperty>
<rdfs:DatatypeProperty rdf:ID="possuiNome">
  <rdfs:domain rdf:resource="#fundador"/>
  <rdfs:range rdf:resource="string"/>
</rdfs:DatatypeProperty>
<rdfs:DatatypeProperty rdf:ID="possuiNome">
  <rdfs:domain rdf:resource="#cidade"/>
  <rdfs:range rdf:resource="string"/>
</rdfs:DatatypeProperty>
```

Fonte: Elaborado pelo autor.

A descrição do RDF baseada em triplas aumenta a expressividade das representações uma vez que os objetos podem ser sujeitos em outros predicados, o que permite a criação de uma rede de significados com semântica explícita. No exemplo anterior, tanto *Cidade* quanto o *Fundador* são sujeitos na relação *possuiNome*, embora apenas uma *Cidade* possua a relação *fundadaPor*. As relações permitidas entre os sujeitos, predicados e objetos são definidas no Esquema RDF ⁶⁷, que permite estabelecer restrições como, por exemplo, impedir declarações semanticamente inválidas como *Fundador* é *fundadoPor* *Cidade*.

No RDF, sujeitos e objetos são valorados a partir da criação de instâncias, por exemplo, *Rio de Janeiro* é instância de *Cidade* e *Estácio de Sá* é instância de *Fundador*.

O trecho de Código fonte 3 apresenta a criação de instâncias para a *Cidade* e *Fundador*.

⁶⁷ Em inglês, *RDF Schema*.

Código fonte 3 – Instâncias da cidade Rio de Janeiro e dados da sua fundação, em RDF

```
<fundador rdf:ID="http://www.exemplo.com.br/modelo/estacio-de-sa">
  <rdfs:label>Estácio de Sá</rdfs:label>
  <possuiNome rdf:datatype="string">Estácio de Sá</possuiNome>
</fundador>

<cidade rdf:ID="http://www.exemplo.com.br/modelo/rio-de-janeiro">
  <rdfs:label>Rio de Janeiro</rdfs:label>
  <possuiNome rdf:datatype="string">Rio de Janeiro</possuiNome>
  <fundadaPor rdf:resource="http://www.exemplo.com.br/modelo/estacio
-de-sa" />
  <fundadaEm rdf:datatype="date">1/3/1565</fundadaEm>
</cidade>
```

Fonte: Elaborado pelo autor.

Com a utilização do RDF já é possível indicar ao computador a semântica dos relacionamentos entre os recursos. No entanto, esta semântica se restringe ao nível do que é explicitamente declarado. Retomando o exemplo de Estácio de Sá e considerando a criação de uma propriedade *nascido em*, é possível declarar que Estácio de Sá é nascido em Portugal. Supondo adicionalmente a intenção de determinar a sua nacionalidade, seria necessário criar outra propriedade *é de nacionalidade* e declarar que Estácio de Sá é de nacionalidade portuguesa. Ao descrever, hipoteticamente, Pedro Álvares Cabral, seria imperioso declarar novamente que é nascido em Portugal e que é de nacionalidade portuguesa, e assim por diante para todas as pessoas e para todas as nacionalidades. A possibilidade de criar uma regra que determina que se uma pessoa nasceu em Portugal ela é de nacionalidade portuguesa, é o que permite o “raciocínio” automático pelo computador, tecnicamente chamado de inferência. Este tipo de descrição é alcançado pelo nível mais alto de descrição da web semântica, a OWL ⁶⁸.

Os três níveis – XML, RDF e OWL – são complementares e compatíveis entre si. Isso quer dizer que os dados descritos sintaticamente em XML podem ser estendidos usando o *framework* RDF de forma a passarem a ter sua semântica descrita de forma explícita. Da mesma forma, estes dados podem receber informações sobre as regras lógicas e passarem a compor a OWL, visando a atender às necessidades de inferência para a integração automática.

⁶⁸ Do inglês *Web Ontology Language*.

A esta estrutura de representação semântica foi dada, na Computação, a inadequada denominação de “ontologia”. Esta afirmação a respeito da inadequação do termo se deve ao fato de “ontologia” ser um termo original da Filosofia (já empregado nesta tese sem necessidade de esclarecimento prévio) e que foi tomado pela Computação para se referir a um tipo de artefato que é, paradoxalmente, puramente pragmático. Para evitar a necessidade de esclarecimento da intenção de uso a cada ocorrência do termo, nesta tese “ontologia” e “ontológico” são empregados estritamente no sentido filosófico. O artefato computacional, doravante será referido, metonimicamente, através do nome da linguagem de programação em que é escrito: OWL ⁶⁹.

Ainda no terreno terminológico, é importante estabelecer claramente os termos utilizados na OWL:

- **Classe (*class*)** diz respeito a uma entidade, por exemplo: `Cidade` e `Fundador`.
- **Instância (*instance*)** diz respeito a um indivíduo específico da entidade, por exemplo: `Rio de Janeiro` e `Estácio de Sá`.
- **Propriedade (*property*)** diz respeito às relações entre as classes. Uma propriedade é dita de domínio quando se refere a um tipo pré definido de dado (por exemplo, a propriedade `possuiNome`, relaciona uma `Cidade` a um domínio *string* ⁷⁰ que significa que uma `Cidade` `possuiNome` que é um texto livre). Quando não é “de domínio”, uma propriedade relaciona uma classe a outra classe (como `Cidade` `fundadaPor` `Fundador`) ou a uma outra propriedade a partir de uma regra lógica (como o caso de `nacionalidade`).

Uma aplicação prática é a possibilidade de exprimir classes genéricas, como, por exemplo, a classe `Pessoa` e declarar `Fundador` como um tipo específico de `Pessoa`. Entretanto, ao contrário das relações hierárquicas dos sistemas convencionais, em um sistema em OWL é possível definir a regra que faz com que uma `Pessoa` seja um `Fundador` (no caso, a regra é fundar uma cidade). Desta forma, é possível dinamizar o modelo, não sendo necessário especificar previamente quem é ou não é um `Fundador`. A partir do momento que uma `Pessoa` é declarada como havendo fundado uma `Cidade`, a regra lógica da OWL “entenderá” que ela é um `Fundador`. Esta regra em OWL é exemplificada no Código fonte 4.

⁶⁹ Exceção feita quando o termo em inglês “*Ontology*” for parte do nome próprio de algum sistema ou projeto, como é o caso da *Music Ontology*, a ser apresentada no capítulo seguinte.

⁷⁰ String é um termo corrente da computação que se refere a um campo de tipo texto (originalmente “*character string*”, em português, “cadeia de caracteres”).

Código fonte 4 – Regra que define um Fundador, expressa em OWL

```
<owl:Class rdf:ID="cidade"/>
<owl:Class rdf:ID="pessoa"/>

<owl:Class rdf:ID="fundador">
  <owl:equivalentClass>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#fundadaPor"/>
      <owl:someValuesFrom rdf:resource="#pessoa"/>
    </owl:Restriction>
  </owl:equivalentClass>
  <rdfs:subClassOf rdf:resource="#pessoa"/>
</owl:Class>

<owl:ObjectProperty rdf:ID="fundadaPor">
  <owl:Domain rdf:resource="#cidade"/>
  <rdfs:range rdf:resource="#pessoa"/>
</owl:ObjectProperty>
```

Fonte: Elaborado pelo autor.

O significativo avanço proporcionado pela OWL para a consecução dos princípios da web semântica consiste, portanto, na flexibilidade trazida pela definição das propriedades. Com elas é possível explicitar para o computador as regras que relacionam as classes e seu funcionamento. De forma nativa, a OWL estabelece alguns tipos de propriedades ⁷¹:

- **Inversa (*Inverse*):** Permite estabelecer uma propriedade correspondente, no sentido inverso do relacionamento. Por exemplo: toda vez que houver um predicado `fundadaPor` por relacionando `Cidade` e `Pessoa` haverá outro predicado `é o fundador de` relacionando `Pessoa` e `Cidade`.
- **Funcional (*Functional*):** Permite declarar que o valor de um predicado para aquele sujeito é único, identificando-o funcionalmente. Não é o caso da relação `possui Nome`, pois pode haver mais de uma cidade com mesmo nome ou mais de uma pessoa com mesmo nome. Nestes casos, é necessário definir identificadores únicos para que o computador seja capaz de lidar com casos de homonímia.
- **Equivalente (*Equivalent*):** Pode se referir tanto a classes, expressando que `Fundador` é equivalente a `Criador`, como pode operar sobre propriedades, permitindo expressar que `é fundador de` equivale a `fundou`. A definição desta

⁷¹ A sumária apresentação a seguir tem por intuito permitir a compreensão do sistema elaborado (descrito no capítulo 6). Para a especificação completa das propriedades, ver a documentação da linguagem OWL. Disponível em: <<https://www.w3.org/TR/owl2-syntax/>>. Acesso em: mai 2016.

propriedade resolve o problema de utilização de metadados diferentes que se referem ao mesmo conteúdo. A propriedade Equivalente aplica-se também para as instâncias, permitindo identificar, por exemplo, que uma instância `Rio de Janeiro` é equivalente a outra denominada `São Sebastião do Rio de Janeiro`.

- **Transitiva (*Transitive*):** Permite declarar que uma propriedade se aplica a todos os componentes de uma hierarquia. Por exemplo: ao definir que `Fundador` é uma `Pessoa` e que `Pessoa` tem a propriedade transitiva `possui Nome`, isso significa, automaticamente, que todo `Fundador` possuirá `nome`.
- **Reflexiva (*Reflexive*):** Permite declarar uma propriedade que age sobre o próprio sujeito. Por exemplo, sabendo-se que `Mem de Sá` “é tio de” `Estácio de Sá`, é possível criar uma propriedade reflexiva `é membro da família`, que permite dizer que `Mem de Sá` e `Estácio de Sá` pertencem a uma mesma família. Declarada como reflexiva, permite afirmar que `Estácio de Sá` é membro da sua própria família.
- **Simétrica (*Symmetric*):** Permite declarar uma propriedade “bidirecional”. Por exemplo, uma definição de propriedade em `Cidade` “faz divisa com”. Se definida como reflexiva, indicaria que se `Rio de Janeiro` faz divisa com `Niterói`, `Niterói` também fará divisa com `Rio de Janeiro`, automaticamente.

As propriedades em OWL tem ainda seus correspondentes de negação: Disjunta (*Disjoint*), Não-Reflexiva (*Irreflexive*) e Assimétrica (*Asymmetric*). As propriedades em OWL podem ser declaradas combinando tipos ou mesmo omitindo-os. Por exemplo, é possível especificar uma propriedade sem determinar se ela é simétrica ou assimétrica. Esta possibilidade de um especificar um “terceiro estado” confere poder expressivo à OWL.

A complexidade e a visão rizomática do conhecimento pressupõem a noção de “sistema aberto”, aquele no qual há interação com o ambiente externo ao qual o sistema se situa. A noção de sistema aberto está diretamente relacionada à capacidade de adaptação e de criatividade (WILDEN, 2013). Do ponto de vista computacional os sistemas são inevitavelmente fechados, pois o computador não é capaz de interagir criativamente nem de ultrapassar os limites daquilo que lhe é declarado. No entanto, a OWL permite, em

comparação a outras abordagens computacionais, “simular” uma aproximação a esta abordagem de sistemas abertos, ao conceber o “terceiro estado” para as propriedades.

Pensada para representar os recursos na web, considerando sua heterogeneidade e multiplicidade de descrições possíveis, a OWL pressupõe uma especificação estrutural (das classes e propriedades básicas) pouco restritiva e concentra as restrições na aplicação de regras lógicas. Esta característica conduz a uma das principais potencialidades da web semântica, que é prover a “interoperabilidade”. Interoperabilidade é a capacidade de um sistema – informatizado ou não – de se comunicar de forma transparente com outro sistema trocando dados de forma integrada e coerente. Um exemplo já apresentado de interoperabilidade é a capacidade de a busca a partir do nome de uma cidade no sistema de pesquisas Google haver buscado informações meteorológicas sobre esta mesma cidade em outro sistema de forma automática.

Dentro do projeto da web semântica, uma das preocupações tem sido precisamente a de garantir a interoperabilidade. O grupo de trabalho LDP – Plataforma de Dados Interligados ⁷² vem conduzindo os estudos da iniciativa denominada Dados Ligados ⁷³ quem tem por objetivo definir padrões aplicáveis e reutilizáveis à descrição dos recursos disponíveis na web de forma a assegurar a correspondência semântica entre recursos de diferentes origens. É esta interligação de recursos que permite, ao sistema Google, ainda considerando o exemplo anterior, obter dados sobre a cidade (descrição e data de fundação) a partir do verbete correspondente na Wikipedia, sem necessidade de duplicação ou cópia destes dados.

Estruturalmente, a OWL permite construir um sistema capaz de abordar a complexidade. O princípio dialógico é alcançável pela possibilidade de definir livremente as propriedades que relacionam as classes. O princípio recursivo é alcançado pela própria natureza das propriedades que podem se referir a outras propriedades e atuar recursivamente na hierarquia de classes. Por fim, o princípio hologramático é alcançável na medida em que cada instância, contem em si a totalidade das regras (propriedades) das classes às quais se submete.

A aproximação da noção computacional de “recurso” com a proposta de “dispositivo infocomunicacional” permite a concepção de um sistema de organização do conhecimento capaz de refletir a “teia de relações” que se estabelecem e podem ser apreendidas dinamicamente. Com esta dinamização é possível ao sistema operar a noção de fronteiras

⁷² Do inglês, *Linked Data Platform*. Disponível em: <<https://www.w3.org/TR/ldp/>>. Acesso em: mai 2016.

⁷³ Do inglês, *Linked Data*.

sobrepostas, como visto, um dos princípios que caracterizam os domínios complexos. No entanto, o outro princípio destes domínios, a gradação das relações, não é plenamente alcançável através das propriedades nativas da OWL. Neste sentido, é necessário estender a OWL, atribuindo-lhe capacidade de expressar e manipular propriedades fuzzy. O projeto “Fuzzy OWL” vem sendo conduzido pelo pesquisador italiano Umberto Straccia e permite criar propriedades que expressam graus de pertinência operando sobre uma mesma relação (STRACCIA, 2012), (STRACCIA, 2013).

Utilizando um exemplo hipotético do domínio musical, em vez de definir uma propriedade *é influenciado por* que só permitiria estabelecer se um determinado músico influencia outro, é possível expressar em graus o quanto um determinado músico influencia outro. Definindo esta propriedade como transitiva, é possível estabelecer desvendar uma complexa rede de influências, exprimível em seu caráter fuzzy. As possibilidades de aplicação destes princípios para um sistema de organização de conhecimento que lide com canções populares são vislumbradas a partir de uma definição dos aspectos metodológicos específicos para este domínio, discutidas no capítulo seguinte.

5 ASPECTOS METODOLÓGICOS RELACIONADOS À CANÇÃO POPULAR

A improvável conjunção de fatores históricos e sociais que formaram as condições de surgimento de uma forma artística única (e em si múltipla) como a canção popular brasileira, é o que decisivamente a caracteriza. O filósofo (e cancionista) Antônio Cícero afirma que a canção popular brasileira procede de modo sintético, pela adição e/ou descarte de elementos que lhe são adventícios, em contraposição, à música clássica que se desdobra de modo mais analítico (como se seus desenvolvimentos fossem uma consequência lógica de seus estágios anteriores) (CICERO, 2005).

Reconhecendo-se, portanto, a canção popular como uma linguagem autônoma e portadora de singularidades definidoras em relação a outras formas de manifestação culturais, faz-se necessário investigar metodologias e ferramentas específicas para sua abordagem. Entretanto, devido ao fato deste reconhecimento ser recente, é inevitável dissociar as iniciativas relacionadas à canção popular de iniciativas de cunho geral, digam elas respeito a outros tipos de música (como a música folclórica ou a música clássica), ou em relação ao suporte (partituras, fonogramas, apresentações etc.) destas manifestações.

Neste capítulo são apresentadas as abordagens metodológicas relacionadas à canção popular, direta ou indiretamente, iniciando pela classificação a partir da noção de gênero musical. Em seguida são apresentadas as técnicas relacionadas ao uso de metadados; as classificações colaborativas e as iniciativas específicas para recursos musicais no ambiente da web. Em seguida são discutidos, a partir de exemplos reais, aspectos metodológicos para a web semântica e a utilização de critérios semióticos e do esquematismo tensivo.

5.1 GÊNERO MUSICAL

A forma prevalente de classificação de objetos musicais é o gênero musical, uma noção herdada da música clássica. Na música clássica, os gêneros se originaram a partir do agrupamento de elementos tais como a estrutura da música, a sua formação instrumental ou a sua função social (por exemplo: sonata, concerto e missa, respectivamente). Na música clássica, os elementos que definem os gêneros são consideravelmente estáveis o que permite que sejam ordenados de forma razoavelmente precisa para definir um gênero musical de maneira coerente (embora na música clássica contemporânea, especialmente a partir do movimento do atonalismo – que se caracteriza pelo rompimento das formas musicais

tradicionais – também a música clássica passou a se defrontar com o problema da classificação em gêneros musicais) (WISNIK, 1989).

Inicialmente, “canção popular” foi mais um item da taxonomia de gêneros utilizada pela música clássica, que já possuía um gênero “canção” para designar as obras vocais clássicas que “musicavam” poesia (SADIE, 1980). No entanto, logo se percebeu que um gênero “canção popular” era insuficiente para comportar a variedade de conteúdos que nele precisavam ser acomodados. Verificou-se que, internamente, além dos componentes musicais (rítmicos, harmônicos e melódicos), a canção popular é constituída pelos assuntos que são abordados nas letras; pelas estruturas narrativas com suas projeções de pessoa, tempo e espaço; pelas nuances interpretativas; pelas variações de significado decorrentes dos arranjos; pelas modalizações envolvidas no discurso; pela sua conjuntura histórica e social etc. E, como já foi discutido, estes elementos não se apresentam isoladamente nem podem ser apartados concorrendo, simultaneamente, para a construção da linguagem sincrética da canção popular (TATIT, 1998).

A impossibilidade de designar um gênero musical às canções populares se reflete pela pluralidade de atribuições desde o início das gravações no Brasil ⁷⁴: “canção”, “cançoneta”, “romança”, “rancho”, “marcha”, “marcha-rancho” (SEVERIANO; MELLO, 1997) e até nas designações irreverentes como as empregadas por Noel Rosa entre as décadas de 1920 e 1930: “colegial”, “faminta”, “fonética”, “maluca”, “humorística” e “anatômica”, por exemplo.

Se nos seus primórdios a canção popular brasileira não demandava uma classificação de gêneros mais rigorosa, com a sua elevação ao posto de representante da cultura nacional a partir dos anos 1930 e, especialmente, a partir do advento das transmissões radiofônicas no Brasil, esta situação mudou. Até aquele momento, o mercado fonográfico brasileiro não contava com a presença das indústrias culturais. Essa transição ocorreu gradualmente até que na década de 1940 pode-se considerar que havia efetivamente uma indústria cultural agindo sobre a gravação e a circulação da música popular no Brasil (ORTIZ, 1988).

Uma vez estabelecida, a indústria cultural aproveitou-se da classificação em gêneros musicais e buscou simplificá-la, redefinindo a taxonomia de gêneros para uma formulação mais limitada e objetiva. O objetivo, no caso, era classificar o público consumidor e não as obras, diminuindo, assim, os riscos de sua atuação comercial (ibidem). Desta redefinição dos gêneros musicais para as canções populares, surgiram denominações sem qualquer

⁷⁴ A primeira gravação musical no Brasil data de 1902. A canção “Isto é Bom”, composição de Xisto Bahia, foi interpretada por Baiano, gravada na Casa Edison, e classificada como “lundu”.

fundamentação musical, como é o caso de MPB – sigla para Música Popular Brasileira – gênero aglutinador surgido na década de 1960 e que abrange qualquer tipo de canção possuidora de características sócio-culturais compatíveis com os perfis de consumo da classe média (NAPOLITANO, 2002).

Com a globalização dos mercados e a difusão das obras musicais pelo mundo, surgiram gêneros como *World Music* e *Latin Music*, que são utilizadas para designar qualquer obra a ser comercializada que esteja fora dos padrões previamente reconhecidos pelos mercados internacionais e que também pouco ou nada informam a respeito das características artísticas dos conteúdos que neles são classificados.

Essa incapacidade das indústrias culturais em classificar as canções populares já havia sido identificada pelos sociólogos da cultura, que desde a década de 1960 afirmam que os padrões de consumo cultural baseados em gêneros e classes não podem ser conformados a partir de teorias preditivas, pois não há uma ordem estável e previsível no nível de produção da cultura popular (WILLIAMS, 1983) e que sempre haverá mais gêneros nos níveis da produção e da oferta do que as teorias são capazes de conceber (DIMAGGIO, 1987). Se esta constatação se aplica à cultura popular internacional, no caso da miscigenada canção popular brasileira, esta impossibilidade é ainda mais acentuada. Na prática, as canções acabam sendo classificadas de acordo com algum gênero musical comercialmente determinado, acompanhada de descrições de conteúdo expressas através de conjuntos de metadados de catalogação.

5.2 METADADOS PARA CANÇÕES POPULARES

Desde o advento das gravações em discos, as canções populares passaram a ser vendidas como produtos. Inicialmente, cada disco continha uma gravação – uma canção.

Com a criação dos discos de 78 rotações por minuto, as canções passaram a ser registradas de forma agrupada. Um disco de 78 RPM usualmente continha duas canções (uma do lado A e outra do lado B). Nos primeiros discos produzidos no Brasil, os “metadados” descritivos eram curiosamente registrados na própria gravação. Antes de o intérprete iniciar o canto, um “narrador” (normalmente o cantor Bahiano ou o diretor da gravadora, João Baptista Gonzaga) anunciava – aos gritos – o título da canção, o gênero, o intérprete e a gravadora, neste caso, a “Casa Edison” (FRANCESCHI, 2002). Por vezes, anunciava também o nome do compositor, o nome do conjunto ou grupo de acompanhantes e o ano. Estes metadados, também sem padronização, eram impressos, constando nas capas e no rótulo dos discos.

No final da década de 1940, com a invenção do *long-play* (LP) a capacidade de armazenamento aumentou e os discos passaram a agrupar mais canções (usualmente entre 10 e 12). Com a invenção do *compact disc* (CD) na década de 1980 houve novo aumento da capacidade de armazenamento. Se inicialmente buscava-se classificar a canção individualmente, com o surgimento dos discos e a atuação da indústria fonográfica buscando vendê-los, passou-se a classificar o agrupamento destas canções (um disco de rock, por exemplo) ou em outros casos, classificava-se o artista (um disco do Altamiro Carrilho ⁷⁵, que é um músico de choro, por exemplo).

Com o avanço da computação e a proliferação de mecanismos de reprodução digital – sendo o mais comum o formato mp3 (*MPEG-1/2 Audio Layer 3*) – a unidade central voltou a ser a canção e não mais o disco. Uma das transformações trazidas pelo formato mp3 é que os próprios usuários passaram a poder descrever os seus arquivos musicais, utilizando o formato “tag ID3”. Na primeira versão da tag ID3, criada em 1996, o conjunto de metadados era composto por 7 campos, conforme listado no Quadro 2:

Quadro 2 – Metadados da tag ID3, versão 1

Atributo	Tamanho	Conteúdo
<i>title</i>	30	título da música
<i>artist</i>	30	nome do artista
<i>album</i>	30	nome do disco
<i>year</i>	4	ano
<i>comment</i>	30	comentário
<i>track</i>	4	número da faixa
<i>genre</i>	1	lista pré definida de 80 gêneros

Fonte: ID3 ⁷⁶

Este limitado conjunto de metadados não permite, por exemplo, especificar o compositor e o intérprete da gravação. Há apenas um genérico campo artista, restrito a 30 caracteres (como todos os demais campos textuais). Este modelo foi adotado como padrão por restringir as descrições a 128 bytes, podendo assim ser incorporado no próprio arquivo de áudio digital (mp3) sem comprometer o seu tamanho nem integridade, sendo até hoje amplamente utilizado pela maioria dos aparelhos de reprodução (*mp3 players*).

Na segunda versão do conjunto de metadados, denominada ID3v2, lançada em 1998, foram definidos 74 campos, cujos metadados são apresentados no Quadro 3:

⁷⁵ Altamiro Aquino Carrilho (1924-2012) foi um compositor e flautista brasileiro.

⁷⁶ Disponível em: < <http://id3.org/>>. Acesso em: mai 2016.

Quadro 3 – Metadados da tag ID3, versão 2

Sigla	Atributo
AENC	<i>Audio encryption</i>
APIC	<i>Attached picture</i>
COMM	<i>Comments</i>
COMR	<i>Commercial frame</i>
ENCR	<i>Encryption method registration</i>
EQUA	<i>Equalization</i>
ETCO	<i>Event timing codes</i>
GEOB	<i>General encapsulated object</i>
GRID	<i>Group identification registration</i>
IPLS	<i>Involved people list</i>
LINK	<i>Linked information</i>
MCDI	<i>Music CD identifier</i>
MLLT	<i>MPEG location lookup table</i>
OWNE	<i>Ownership frame</i>
PRIV	<i>Private frame</i>
PCNT	<i>Play counter</i>
POPM	<i>Popularimeter</i>
POSS	<i>Position synchronisation frame</i>
RBUF	<i>Recommended buffer size</i>
RVAD	<i>Relative volume adjustment</i>
RVRB	<i>Reverb</i>
SYLT	<i>Synchronized lyric/text</i>
SYTC	<i>Synchronized tempo codes</i>
TALB	<i>Album/Movie/Show title</i>
TBPM	<i>BPM (beats per minute)</i>
TCOM	<i>Composer</i>
TCON	<i>Content type</i>
TCOP	<i>Copyright message</i>
TDAT	<i>Date</i>
TDLY	<i>Playlist delay</i>
TENC	<i>Encoded by</i>
TEXT	<i>Lyricist/Text writer</i>
TFLT	<i>File type</i>
TIME	<i>Time</i>
TIT1	<i>Content group description</i>
TIT2	<i>Title/songname/content description</i>
TIT3	<i>Subtitle/Description refinement</i>
TKEY	<i>Initial key</i>
TLAN	<i>Language(s)</i>
TLEN	<i>Length</i>
TMED	<i>Media type</i>
TOAL	<i>Original album/movie/show title</i>
TOFN	<i>Original filename</i>
TOLY	<i>Original lyricist(s)/text writer(s)</i>
TOPE	<i>Original artist(s)/performer(s)</i>
TORY	<i>Original release year</i>
TOWN	<i>File owner/licensee</i>
TPE1	<i>Lead performer(s)/Soloist(s)</i>
TPE2	<i>Band/orchestra/accompaniment</i>
TPE3	<i>Conductor/performer refinement</i>
TPE4	<i>Interpreted, remixed, or otherwise modified by</i>
TPOS	<i>Part of a set</i>
TPUB	<i>Publisher</i>

TRCK	<i>Track number/Position in set</i>
TRDA	<i>Recording dates</i>
TRSN	<i>Internet radio station name</i>
TRSO	<i>Internet radio station owner</i>
TSIZ	<i>Size</i>
TSRC	<i>ISRC (international standard recording code)</i>
TSSE	<i>Software/Hardware and settings used for encoding</i>
TYER	<i>Year</i>
TXXX	<i>User defined text information frame</i>
UFID	<i>Unique file identifier</i>
USER	<i>Terms of use</i>
USLT	<i>Unsynchronized lyric/text transcription</i>
WCOM	<i>Commercial information</i>
WCOP	<i>Copyright/Legal information</i>
WOAF	<i>Official audio file webpage</i>
WOAR	<i>Official artist/performer webpage</i>
WOAS	<i>Official audio source webpage</i>
WORS	<i>Official internet radio station homepage</i>
WPAY	<i>Payment</i>
WPUB	<i>Publishers official webpage</i>
WXXX	<i>User defined URL link frame</i>

Fonte: ID3

Apesar de extenso, o conjunto ID3 versão 2 ainda não é suficiente para exprimir os dados específicos das canções populares e, como pode ser percebido a partir da leitura do quadro anterior, tampouco apresenta praticidade de uso, incorrendo nos problemas de padrões de metadados discutidos no capítulo anterior, quais sejam: a falta de padronização no preenchimento; a ocorrência de dados incompletos ou não preenchidos; e, sobretudo, a falta de semântica.

Estudos voltados para os repertórios brasileiros vem explorando as dificuldades de descrição de conteúdos musicais utilizando padrões de metadados, sejam estes específicos, como os definidos pela *Library of Congress*⁷⁷, sejam os de propósito geral, como é o caso do padrão Dublin Core⁷⁸ (STANECK; LANZELOTTE, 2007), (DE SANTIS; LANZELOTTE, 2014).

Com o advento da web no final do século XX e as transformações ocorridas na produção e no consumo de bens culturais, o primeiro mercado a refletir as alterações foi o das canções populares. Com o declínio das vendas de discos, as indústrias fonográficas começaram a abandonar a concepção da música como produto para tratá-la como informação

⁷⁷ A *Library of Congress* define mais de uma dezena de padrões de metadados. Para informações a respeito de cada um, ver: <<https://www.loc.gov/standards/>>. Acesso em: mai 2016.

⁷⁸ Para mais detalhes, ver a especificação do padrão. Disponível em: <<http://dublincore.org/>>. Acesso em: mai 2016.

e serviço (DIMAGGIO *et al.*, 2001). Esta mudança acarretou a volta do investimento por parte das indústrias culturais na construção de sistemas para a recomendação de canções e para a venda sob demanda de arquivos musicais digitais através de aplicativos e serviços de rádios online (ADOMAVICIUS; TUZHILIN, 2005).

Com esta mudança de paradigma que afeta a produção e o consumo de música, somados às implicações da evolução tecnológica, novas possibilidades comerciais vem sendo desenvolvidas e novas formas de classificação das canções populares começaram a surgir. O compartilhamento dos arquivos mp3 e a posterior construção de aplicativos que buscavam e indexavam as canções a partir das *tags* ID3 foi uma das primeiras iniciativas de métodos colaborativos.

5.3 MÉTODOS COLABORATIVOS

Com os recursos disponíveis na web social ou web colaborativa, os utilizadores de sites de música passaram a poder aplicar *tags* personalizadas diretamente nos recursos compartilhados na web, e não mais em seus arquivos pessoais, constituindo, assim, bases globais de dados músicas descritas colaborativamente.

Estudos dedicados à análise do uso de *tags* em sistemas colaborativos de classificação de música apontam que a *collaborative tagging*⁷⁹ revela categorias que não são contempladas pelos metadados convencionais, como a classificação por humor (se a canção é calma ou agitada, triste ou alegre), por opiniões pessoais (legal, chato, gostei, não gostei) ou por locais ou situações adequados para se ouvir a canção (festa, jantar, elevador, corrida), por exemplo. Por outro lado, estes estudos também revelam que o número de *tags* criadas pelos usuários tende ao infinito, uma vez que a não adoção de um vocabulário controlado acaba multiplicando indefinidamente a atribuição de novas *tags*. Outro aspecto revelado aponta que, embora a quantidade de *tags* cresça indefinidamente, há uma tendência de concentração do uso de um conjunto menor de *tags* para um conjunto menor de objetos. Na prática, as canções mais conhecidas recebem muitas classificações iguais enquanto a maior parte das canções fica sem qualquer classificação, o que acarreta desvios e uma tendência à massificação nas recomendações feitas por sistemas computacionais (THOMPSON, 2008) e (SANTINI, 2010).

⁷⁹ Preferiu-se manter a expressão em inglês, por não haver uma tradução consensual.

Com a diminuição do entusiasmo em relação às abordagens colaborativas, novas ferramentas e metodologias vem sendo desenvolvidas para lidar com objetos musicais na web.

5.4 ABORDAGEM NA WEB

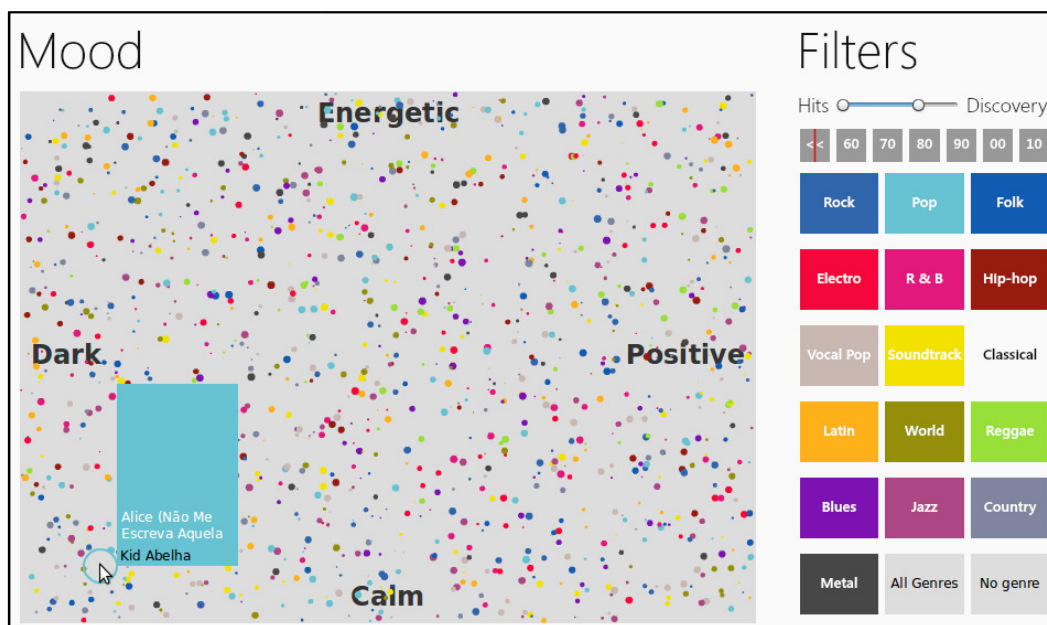
Considerando a grande quantidade de sites e sistemas dedicados à música atualmente disponíveis na web, seria inviável realizar uma análise completa e imparcial. São apresentados, portanto, três trabalhos considerados originais e que permitem, de forma complementar, verificar diferentes visões: MusicOvery, Aquela Música e Musica Brasilis.

Musicoverly

O site francês Musicoverly⁸⁰, é uma sistema de rádio online que, simultaneamente aos metadados tradicionais (título, artista, ano e gênero musical) possibilita aos utilizadores escolherem músicas a partir de uma matriz de “humores”, situando cada música como um ponto em um espaço bidimensional, no qual o eixo vertical contem nas extremidades os valores *calm* e *energetic*, e o eixo horizontal contem nas extremidades os valores *dark* e *positive*, conforme mostra a Figura 16.

⁸⁰ Disponível em: <<http://www.musicoverly.com>>. Acesso em: mai 2016.

Figura 16 – Tela do site Musicoverly



Fonte: Adaptado do site Musicoverly.

A possibilidade de expressar a gradação, refletindo, a partir do posicionamento no espaço bidimensional, a intensidade dentro de categorias predeterminadas constitui um avanço significativo no sentido de substituir os tipos discretos e mutuamente exclusivos dos sistemas tradicionais. Entretanto, o site Musicoverly baliza as categorizações nos resultados obtidos através de algoritmos de análise de frequências sonoras. A comparação dos resultados obtidos serve para determinar o “humor” das músicas (PLEWA; KOSTEK, 2013).

As análises computacionais automatizadas, sejam baseadas em conteúdo ou em abordagens estatísticas, vem apresentando melhoras significativas em seus resultados, especialmente em decorrência do trabalho da sociedade ISMIR – *International Society of Music Information Retrieval* – fundada em 2000 e que conduz as principais pesquisas na área (DOWNIE; BYRD; CRAWFORD, 2009). No entanto, claramente os resultados obtidos por este tipo de abordagem (ainda) não são fidedignos, como revelado na Figura 19, na qual um exemplo não intencional de utilização encontrou a canção brasileira Alice (Não Me Escreva Aquela Carta de Amor), do grupo Kid Abelha como uma canção *calm* e *dark*, embora presumivelmente qualquer categorização manual feita por brasileiros fosse situá-la justamente na outra extremidade.

Aquela Música

Em um trabalho anterior, o projeto “Aquela Música”⁸¹, o autor da presente tese apresentou a construção de uma aplicação baseada na análise de canções populares a partir de seus fragmentos (DE SANTIS, 2010), (DE SANTIS, 2013), conforme ilustra a Figura 17.

Figura 17 – Tela do site Aquela Música



Fonte: Adaptado do site Aquela Música.

O projeto “Aquela Música” considerava que, partindo dos fragmentos de dados existentes nas canções (aspectos melódicos, harmônicos, rítmicos, da letra e do discurso) e através de uma navegação multidimensional a partir destes fragmentos seria possível representar e recuperar dados de canções populares. O projeto recebeu este nome por causa do seu intuito inicial que era o de responder à pergunta feita corriqueiramente quando se tenta lembrar ou descobrir uma música da qual se tem poucas ou fragmentadas informações: “Como é mesmo aquela música? Aquela que fala de um dado assunto, aquela que foi

⁸¹ Disponível em: <<http://www.aquelamusica.com.br>>. Acesso em: mai 2016.

interpretada por determinado cantor ou que tem determinado músico tocando flauta” (ibidem). Embora tenha alcançado o intuito inicial ao permitir a navegação multidimensional, o fato de ignorar que a complexidade da canção não é alcançável pela “soma das partes”, o caminho de reconstrução das canções a partir dos fragmentos mostrou-se falacioso e operacionalmente pouco promissor ⁸².

Musica Brasilis

O portal Musica Brasilis ⁸³, criado em 2009 pela pesquisadora, cravista e doutora em Informática, Rosana Lanzelotte tem por objetivo o resgate e a difusão de repertórios brasileiros de todos os tempos. Embora preponderantemente voltado para a música grafada (partituras musicais), devido à preocupação que motivou a sua criação (a inacessibilidade de edições voltadas à execução), o site oferece objetos musicais em diferentes formatos (vídeos, textos, jogos e escutas guiadas).

Concebido desde o início para o ambiente da web, o projeto se preocupa com a interoperabilidade dos conteúdos e em não se constituir como mais uma iniciativa centralizadora, das quais Lanzelotte destaca o caso do RISM – Repertório Internacional de Fontes Musicais ⁸⁴ – criado em 1952 com o intuito de concentrar, em uma base única, metadados sobre manuscritos de instituições espalhadas pelo mundo e que mantem a mesma proposta até os dias atuais (DE SANTIS; LANZELOTTE, 2014). Na busca da interoperabilidade, todos os conteúdos do portal Musica Brasilis são identificados através de URIs, os vídeos de todas as gravações estão publicados na plataforma Youtube e integrados ao portal. Além disso, o portal promove trocas de dados com outras iniciativas similares, como o projeto SESC partituras ⁸⁵.

Segundo Lanzelotte, há a previsão de aplicar uma camada de descrição em RDF sobre os recursos musicais do portal (ibidem, p.320), configurando, assim, o projeto Musica Brasilis como a primeira iniciativa brasileira do domínio musical alinhada à web semântica, que se tem notícia.

⁸² Apesar de atualmente considerado superado, o projeto cumpriu seu papel e obteve relativo êxito, recebendo inclusive o Prêmio Funarte de Reflexão Crítica e Produção Cultural para Internet, oferecido pelo Ministério da Cultura no ano de 2010.

⁸³ Disponível em: <<http://www.musicabrasilis.org.br>>. Acesso em: mai 2016.

⁸⁴ Do original, em francês, *Répertoire International des Sources Musicales*

⁸⁵ Disponível em: <<http://www.sesc.com.br/SescPartituras/>>. Acesso em: mai 2016.

5.5 ABORDAGEM NA WEB SEMÂNTICA

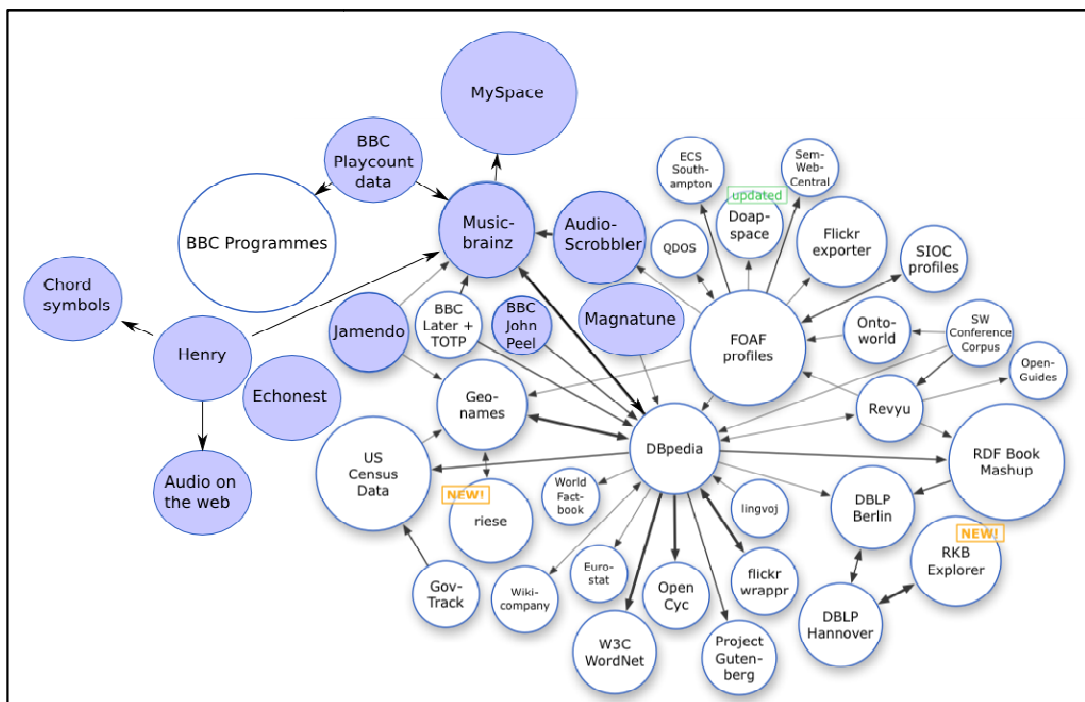
Algumas abordagens a sistemas musicais vem surgindo no âmbito da web semântica. Destas, foram selecionadas duas para discussão de suas características, limitações e potencialidades: DBTune e Linked Jazz.

DBTune

No momento de realização desta pesquisa, o DBTune ⁸⁶ é a única iniciativa relacionada ao domínio musical na web semântica vinculada ao projeto LDP, da W3C.

A proposta do DBTune é mapear semanticamente os registros de fonogramas e álbuns das coleções de diversos sítios participantes da iniciativa, como MySpace, MusicBrainz e a rádio da BBC. Em maio de 2016, o projeto contabilizava mais de 14 bilhões de triplas RDF relacionadas ao domínio musical. Os conjuntos de dados ⁸⁷ e suas relações estão representados pelo grafo da Figura 18.

Figura 18 – Grafo dos conjuntos de dados do projeto DBTune



Fonte: Adaptado do site DBTune.

⁸⁶ Disponível em: <<http://www.dbtune.org>>. Acesso em: mai 2016.

⁸⁷ Do original, em inglês *datasets*.

Como é característico dos projetos na web semântica, os conjuntos de dados originais de cada site ou sistema são mantidos em seu formato original, recebendo uma camada de descrição semântica. Esta descrição é feita a partir de diferentes esquemas, compatíveis e integrados entre si. Um dos principais esquemas de descrição voltados para o domínio musical, e utilizado pelo DBTune, é a “Music Ontology” (RAIMOND *et al.*, 2007).

A Music Ontology foi desenvolvida especificamente para a descrição de fonogramas, artistas e álbuns. Escrita em RDF e OWL, a Music Ontology contém um conjunto de 54 classes e 153 propriedades ⁸⁸. Conforme mencionado anteriormente, uma prerrogativa importante dos modelos OWL é permitir a integração com outros esquemas. No caso da Music Ontology, as classes e propriedades relacionadas a artistas remetem ao projeto FOAF ⁸⁹, enquanto as informações descrições biográficas remetem ao projeto DBpedia. Isto proporciona a interoperabilidade de dados e diminui as redundâncias de dados, tão comuns na web.

No entanto, devido à inerente característica de instabilidade dos recursos digitais na web (DAY, 2006) e também à incipiência das tecnologias da web semântica, muitos destes projetos sofrem com a recorrente indisponibilidade de seus recursos, causadas por mudanças de endereço, exclusão de conteúdo ou mesmo mudança de políticas de acesso. Durante o curso desta pesquisa, estes tipos de ocorrência foram frequentes:

- Durante o ano de 2016, por diversas vezes tentou-se utilizar as ferramentas disponíveis no site do projeto DBTune e as mesmas estavam indisponíveis.
- O acesso aos recursos do MySpace se tornaram permanentemente indisponíveis devido a alterações de permissão de acesso e modificação dos URIs.
- O projeto Music Ontology deixou de ser “código aberto” ⁹⁰ e passou a ser controlado por um grupo restrito de pesquisadores. Isto acarreta que erros estruturais e insuficiências do modelo não podem ser corrigidos pela comunidade de utilizadores, o que leva a W3C a classificá-la com restrições na lista de “boas ontologias” ⁹¹.

⁸⁸ Especificação disponível em: < <http://musicontology.com/specification/> >. Acesso em: mai 2016.

⁸⁹ Acrônimo para *Friend of a Friend*. Disponível em < <http://www.foaf-project.org/> >. Acesso em: mai 2016.

⁹⁰ Do original, em inglês *open source*. Tipo de sistema ou programa computacional em que o código fonte é tornado disponível para a comunidade de utilizadores, outorgando-lhe o direito de consultar, modificar e distribuir o sistema ou programa gratuitamente.

⁹¹ Do original, em inglês, *Good Ontologies*. Disponível em <https://www.w3.org/wiki/Good_Ontologies#The_Music_Ontology>. Acesso em: mai 2016.

Estas dificuldades tem levado ao surgimento de iniciativas conduzidas por empresas ou instituições de pesquisa e que transcorrem fora dos grupos de trabalho da W3C. Estas iniciativas vem apresentando resultados significativos, como é o caso de um projeto relacionado à música, o *Linked Jazz*.

Linked Jazz

O *Linked Jazz* é um projeto dirigido por Cristina Pattuelli, do *Pratt Institute School of Information*, do programa *Library and Information Science*, em Nova Iorque. Este projeto realiza a construção de uma camada semântica em OWL baseada na análise dos dados disponíveis na DBPedia, feitas especificamente para o domínio do jazz (PATTUELLI; WELLER; SZABLYA, 2011), (PATTUELLI; RUBINOW, 2013).

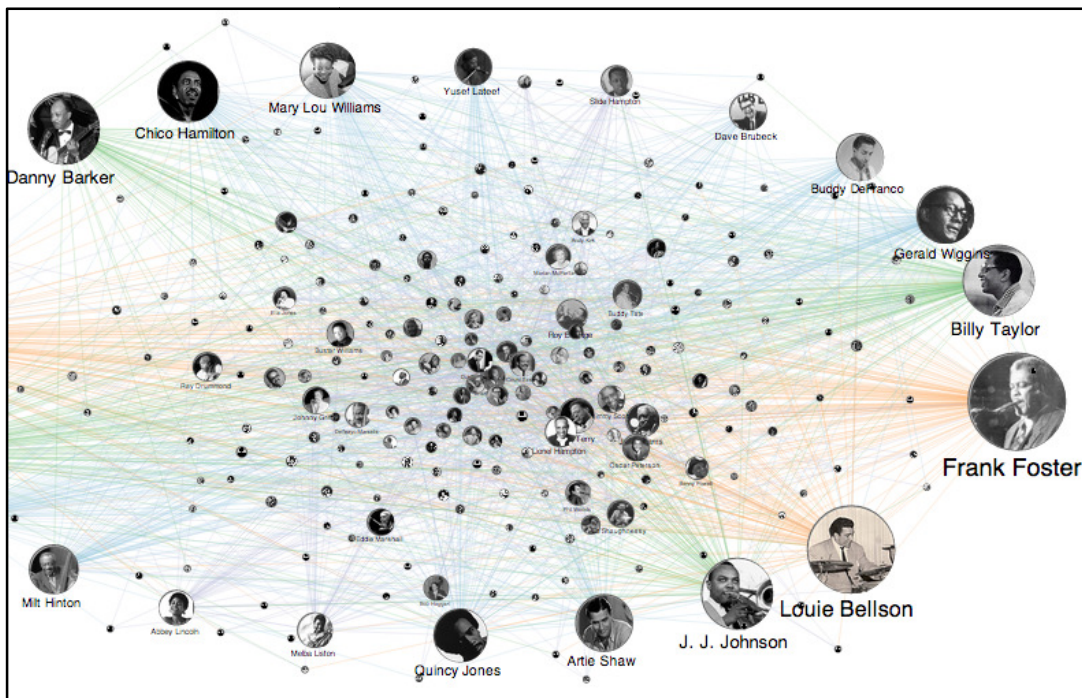
O projeto mapeia os artistas e as suas relações consideradas mais relevantes para o domínio do jazz, a saber:

- “encontrou”
- “é conhecido de”
- “é amigo de”
- “é amigo íntimo de”
- “é influenciado por”
- “é o guru de”
- “colaborou com”
- “esteve em uma banda com”
- “tocou junto com”
- “formou uma banda com”
- “participou de turnê com”
- “foi líder da banda de”⁹²

⁹² Do original, em inglês: *"has met"* ; *"is an acquaintance of"* ; *"is a friend of"* ; *"is a close friend of"* ; *"is influenced by"* ; *"is a mentor of"* ; *"collaborated with:"* ; *"was in a band together with"* ; *"played with"* ; *"was a member of the band of"* ; *"toured with"* ; *"was the bandleader for"*. Disponível em: <<https://linkedjazz.org/about-the-project/>>. Acesso em: mai 2016.

A Figura 19 apresenta a visão da tela de navegação do *Linked Jazz*.

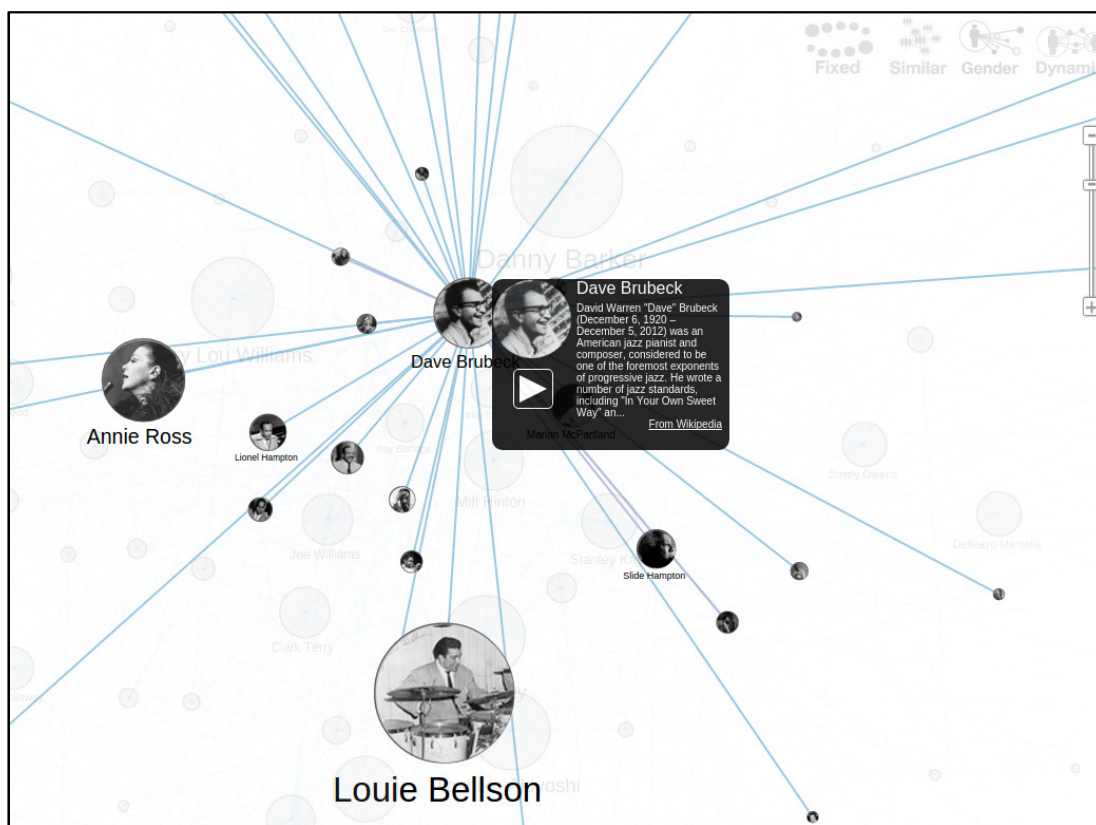
Figura 19 – Tela de navegação do projeto *Linked Jazz*



Fonte: Adaptado do site *Linked Jazz*.

Ao selecionar um determinado artista, são exibidas as suas relações. O sistema *Linked Jazz* comprova a possibilidade de representação visual de uma organização do conhecimento baseada na noção de rede, estabelecendo cada artista como um nó, e permitindo, a partir de suas relações, a conexão com os demais nós. Além disso, o projeto utiliza os recursos de integração de dados providos pela web semântica. A Figura 20 ilustra a seleção de um nó da rede, o artista Dave Brubeck (1920-2012) e sua biografia resumida obtida a partir da integração com os dados da DBPedia.

Figura 20 – Tela do projeto *Linked Jazz*. Seleção de um artista



Fonte: Adaptado do site *Linked Jazz*.

Este tipo de iniciativa ilustra positivamente as possibilidades trazidas pelas tecnologias relacionadas à web semântica, ao mesmo tempo em que revela a necessidade de um trabalho amplo e de longa duração. Iniciado em 2011, o projeto envolve uma dezena de pesquisadores e desenvolveu uma série de ferramentas específicas, como “*The Linked Jazz Transcript Analyzer*”, “*Name Mapping and Curator Tool*”, “*Ecco! a Linked Open Data (LOD) application for entity resolution*”, “*Linked Jazz 52nd Street*” e “*Linked Jazz Network Visualization Tool*”⁹³.

Em suma, iniciativas como as que estão em curso na web e na web semântica, trazem novas formas de abordar a música para além do gênero musical e dos metadados tradicionais, revelando a aplicabilidade prática das inovações tecnológicas. Porém, no que concerne ao domínio abordado nesta tese, a dificuldade recai novamente na falta de critérios específicos para a canção popular brasileira.

⁹³ Informações obtidas no site do projeto. Disponível em: <<http://www.linkedjazz.org>>. Acesso em: mai 2016.

As iniciativas que abordam aspectos fuzzy do domínio musical vem sendo desenvolvidas para repertórios musicais genéricos e, em sua aparente totalidade, a partir de análises baseadas em algoritmos de leitura de frequências de ondas sonoras (YANG; LIU; CHEN, 2006), (GUAN; YUEN; COENEN, 2015) . Este tipo de abordagem não permite captar as gradações, as sobreposições e, sobretudo, as relações que a música estabelece interna e externamente, necessárias para considerar a música como um dispositivo e não como um objeto autocontido, fechado sobre si mesmo. Por outro lado, estudos com estes objetivos vem sendo desenvolvidos há várias décadas pela Musicologia e recebendo contribuições de estudos realizados no campo da Semiótica Musical.

5.6 ABORDAGEM SEMIÓTICA / TENSIVA

Estudos da Musicologia tem investigado caminhos para lidar com as particularidades do objeto musical, sendo a Semiótica uma das abordagens adotadas com sucesso. O musicólogo Jean-Jacques-Nattiez afirma que para abordar semioticamente ⁹⁴ a música primeiramente “é preciso admitir que a música é um fenômeno simbólico, e que a obra musical constitui, portanto, uma forma simbólica” (NATTIEZ, 2003, p. 12). Embora se dedique especialmente à música clássica escrita, abordando a representação através de partituras musicais, Nattiez afirma que “da mesma forma, uma gravação também deixa de conter apenas ondas sonoras para representar o objeto sonoro correspondente a uma obra musical” (ibidem, p.12).

As análises semióticas realizadas por Nattiez tomam por base princípios semióticos estabelecidos pelo pesquisador francês Jean Molino, e seu construto teórico de “fato musical” a partir de um modelo tripartite, e suas dimensões poiética, estética e neutra.

Para Molino (1975) todas as ações humanas deixam vestígios materiais, portanto, também a música tem uma realidade material, que constitui formas simbólicas, uma vez que são portadoras de significações tanto para aqueles que a produzem como aqueles que a percebem. Segundo Molino, mesmo que a música seja destituída de qualquer significação intencional, uma forma simbólica sempre resulta de um processo criador passível de ser descrito ou reconstituído. A esta dimensão ele chama de poiética. A segunda dimensão,

⁹⁴ Nattiez utiliza o termo “semiologia” em vez de “semiótica”. Para fins de simplificação terminológica, considerando que não há prejuízo semântico nem ofensa às ideias do autor, será adotado apenas o termo “semiótica” e seus derivados.

denominada estésica, e corresponde às múltiplas redes de significações que os receptores de uma obra musical atribuem à forma simbólica. Molino destaca, no entanto, que os receptores não “recebem” a significação, mas sim a constroem em um processo ativo de percepção a partir dos vestígios materiais. O que viabiliza este processo é precisamente o vestígio, uma vez que (na maioria dos casos) a dimensão poiética não é diretamente inteligível. Molino denomina este vestígio de dimensão neutra. Nattiez, no entanto considera o termo polêmico e prefere designá-lo como “dimensão material” ou “dimensão imanente”.

Nattiez define a sua semiótica musical “não como uma ciência da comunicação, mas sim como um estudo da especificidade do funcionamento das formas simbólicas” (NATTIEZ, 2003, p. 15), ou seja, da análise da organização material da dimensão imanente e dos fenômenos de remissão provocados por estes signos. Remissões essas provocadas a partir dos vestígios e que se repartem entre o polo poiético e o polo estésico. Nattiez define ainda que a análise da dimensão imanente descreve formas e estruturas mais ou menos regulares (segundo ele, apesar de considerar o termo pedante, seria preferível falar em “quase-estruturas”). Já as análises poiéticas e estésicas, buscam descrever e interpretar os processos (ibidem).

É, portanto, a identificação das regras analíticas que operam as dimensões poiética e estésica que permite uma abordagem sistemática ao todo complexo. O nível imanente de Nattiez condiz com a noção de documento (no caso específico desta tese, o documento que corresponde a uma gravação) e a partir do qual é possível retrair o duplo devir deleuzeano, com as linhas de visibilidade da enunciação e as linhas de fissura do dispositivo infocomunicacional.

As aplicações deste modelo semiótico vem apresentando relevantes resultados no âmbito da música clássica (DELALANDE, 1998), (SPOLADORE, 2009), (SAMPAIO, 2014). No que tange à canção popular, desde o reconhecimento da sua linguagem autônoma, originada da entoação da voz, o pesquisador brasileiro Luiz Tatit, vem conduzindo as pesquisas do que se denominou Semiótica da Canção (TATIT, 1994), a partir de estudos da semiótica francesa, especialmente dos trabalhos de Greimas e Zilberberg e do linguista dinamarquês Louis Hjelmslev (1899-1965).

Ao desenvolver uma metodologia analítica específica para o texto sincrético da canção, Tatit estabelece princípios isomórficos entre letra e melodia, correlacionando-os ao plano da expressão (melódico) e plano do conteúdo (letra) (ibidem). A metodologia analítica proposta por Tatit é de caráter preponderantemente discursivo, cujo produto relaciona-se à identificação e à sistematização dos efeitos de sentido do discurso cancional. Na Linguística contemporânea, é reconhecido que o sentido não se constitui apenas pelo reconhecimento dos

enunciados de uma língua expressos no nível imanente, pois a língua não é um código a ser decifrado. Da mesma forma, o sentido também não é individualmente determinado pelo compositor nem pelo ouvinte nos níveis poético e estético, respectivamente, pois é necessário que os enunciados estejam associados aos discursos, cuja natureza é sempre social. Por este motivo, considera-se que, em vez de significados, os discursos produzem efeitos de sentido, constituídos nas relações entre os três níveis frente às condições sociais de produção do enunciado (FIORIN, 1996).

Os perfis analíticos propostos por Tatit, portanto, não são discretas nem dicotômicas, mas sim baseadas nos princípios do esquematismo tensivo sendo, conseqüentemente, mensuráveis em graus e passíveis de sobreposição.

A operacionalização do modelo da Semiótica da Canção se dá a partir da medição da duração e da intensidade, obtidas a partir de grandezas mensuráveis do discurso musical, como o andamento e a tessitura; e do discurso linguístico, como a silabação e a altura ⁹⁵. A medição da intensidade refere-se às variações de frequência dos sons musicais e articula-se em expansão x concentração. A duração refere-se diretamente ao tempo e articula-se em aceleração x desaceleração.

Das correlações entre aceleração x desaceleração e expansão x concentração, em seus variados graus de mistura, resultam os perfis melódicos das canções populares, que Tatit sintetiza em três grandes grupos: “passionalização”, “tematização” e “figurativização”, resumidos a seguir e representados graficamente na Figura 21 (TATIT, 1998).

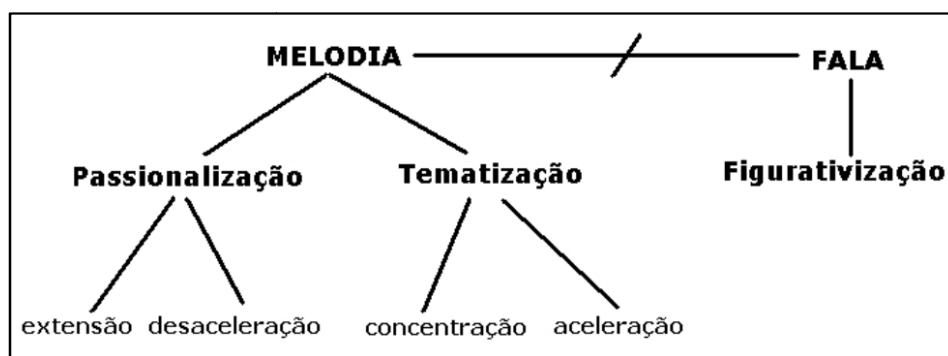
- A passionalização corresponde a melodias “verticalizadas”, com grande exploração da tessitura e um andamento desacelerado (grandes saltos intervalares e ênfase nas durações), que no plano do conteúdo normalmente está associada à idéia de sofrimento, de perda e, conseqüentemente, de busca, de progressão e de surpresa, denotando sentimento de ausência, onde o sujeito evidencia a falta do objeto de desejo.
- A tematização corresponde a melodias “horizontalizadas”, com pequena exploração da altura e grande repetição de motivos de curta duração (concentração da tessitura e aceleração no andamento). As canções deste tipo

⁹⁵ Para definições mais precisas acerca das grandezas musicais, ver, DOURADO, Henrique Autran. *Dicionário de termos e expressões da música*. São Paulo: Editora 34, 2004. No entanto, considera-se que o entendimento de senso comum dos termos empregados (andamento, altura, intensidade) é suficiente. O menos corrente é tessitura, que refere-se à amplitude, à diferença entre a nota mais aguda e a nota mais grave.

são fortemente caracterizadas pela reiteração, que caracterizam sua tendência involutiva, como um eterno retorno ao ponto de partida.

- A figurativização se apresenta como a voz que fala dentro da melodia, em um plano da expressão onde irrompem traços da fala desestabilizando os perfis melódicos temáticos e passionais.

Figura 21 – Tipologia das canções segundo a Semiótica da Canção



Fonte: Adaptada de (DIETRICH, 2007).

O modelo de análise da Semiótica da Canção vem sendo aplicado em análise de canções populares (CARMO JR., 2005), (COELHO, 2003), (DIETRICH, 2003), (TATIT, 2001). Um destes pesquisadores, José Roberto do Carmo Jr., formulou a noção de “coeficiente tensivo” (CARMO JR., 2014) que permite medir a “expressão de primitivos emocionais” básicos, como “alegria”, “tristeza”, “medo” etc. a partir de fundamentos da canção, por ele assim denominados:

- Coeficiente cinemático (Cc), medida da velocidade com a qual é executada uma melodia, tomada como o número de sílabas por minuto;
- Coeficiente entoacional (Ce), extensão tonal da melodia de uma canção, ou seja, a diferença em semitons entre a nota mais grave e a mais aguda;
- Coeficiente modal (Cm), de acordo com o modo (maior ou menor);
- Coeficiente dinâmico (Cd), o correlato acústico da ativação, medido em sone (μS).

Embora não mencione o site francês, em seus resultados preliminares, Carmo Jr. constroi uma abordagem próxima àquela do Musicoverly, porém adequada à canção popular.

Este heterogêneo referencial teórico e metodológico reflete o que se anunciou desde o início como o grande desafio da pesquisa em domínios complexos: o de não incorrer na tentação da ciência simplificadora de “resolver” o problema ou buscar uma solução única e definitiva. A partir de todo este referencial, que certamente, não é exaustivo, buscou-se identificar as possibilidades de abordagem ao domínio da canção popular brasileira, cuja aplicação e resultados são discutidos no capítulo seguinte.

6 CAMPO EMPÍRICO – APLICAÇÃO E RESULTADOS

Neste capítulo é apresentada a aplicação do sistema de organização do conhecimento baseado nos princípios do método complexo e do esquematismo tensivo ao domínio das canções populares brasileiras. Este sistema foi construído para o paradigma da web semântica, dirigido pela recuperação, escrito em linguagem OWL e utilizando propriedades fuzzy. As canções são tomadas como dispositivos infocomunicacionais, atualizadas através do documento correspondente a uma gravação, a qual é analisada com base nos princípios da teoria semiótica.

Inicialmente, é apresentado o corpus analítico e os critérios utilizados para sua determinação.

6.1 *CORPORA DE ANÁLISE*

Para este trabalho foi delimitado um corpus geral, contendo um conjunto abrangente de canções populares e um corpus analítico, de tamanho reduzido, que foi sendo construído empiricamente a partir das relações de uma canção específica, escolhida arbitrariamente.

A tarefa de delimitar o corpus geral selecionando um subconjunto dentro do imenso repertório de canções populares brasileiras foi facilitada pela pesquisa dos musicólogos Jairo Severiano e Zuza Homem de Mello. Em sua clássica obra “A canção no tempo”⁹⁶, os pesquisadores chegaram a um total de quase três mil canções, das quais elegeram, segundo critérios qualitativos, de repercussão popular e através de levantamentos com especialistas, 869 gravações como sendo as mais representativas (SEVERIANO; MELLO, 1997, p. 9–11). Este corpus completo foi utilizado para testes com volume de dados e para verificação dos resultados de navegação a partir das restrições com os dados de catalogação mais imediatos, a saber: título, título alternativo, compositores, intérpretes, ano da gravação e gênero musical atribuído.

Entretanto, de acordo com o intuito de propor um sistema de organização do conhecimento que aborde a canção como objeto complexo, considerando suas fronteiras sobrepostas e as redes de relações mensuráveis em graus, a atividade de analisar e descrever

⁹⁶ Na ocasião do lançamento de uma nova edição em janeiro de 2016, o livro foi saudado como “A Bíblia” da canção popular brasileira. Ver artigo de João Máximo, publicado no Jornal O Globo em 27 de dezembro de 2015. Versão digital disponível em: <<http://oglobo.globo.com/cultura/musica/biblia-musical-do-pais-cancao-no-tempo-ganha-nova-edicao-18368250>>. Acesso em: abr 2016.

todas as canções no âmbito desta tese seria potencialmente irrealizável e de pouco interesse prático, uma vez que não haveria sentido em operar o modelo com um corpus tão estendido sem um critério que estabeleça um ponto de partida. Seguindo o princípio hologramático do método complexo, é plausível considerar a possibilidade de escolher qualquer canção como ponto de partida, uma vez que, partindo de qualquer uma delas é possível “puxar o fio da meada” que permite explorar o tecido complexo do domínio das canções populares brasileiras. Desta forma, optou-se por escolher aleatoriamente uma canção dentre as identificadas por Severiano e Melo. A escolha foi realizada de forma aleatória: com todas as 869 gravações disponíveis em formato mp3, foi criada uma lista de execução em um aplicativo tocador de música (*mp3 player*) e foi acionado o modo *shuffle* (embaralhar). A primeira canção que tocou foi Querelas do Brasil, escolhida, de forma estocástica, como ponto de partida.

6.2 ANÁLISE A PARTIR DE QUERELAS DO BRASIL

O corpus analítico foi construído segundo as regras para abordagem a domínios complexos anteriormente enunciadas, ou seja, a partir das redes de relações obtidas a partir do dispositivo infocomunicacional selecionado, neste caso, a canção Querelas do Brasil.

O primeiro passo consistiu na identificação dos atributos fundamentais do dispositivo, apresentados no Quadro 4, utilizando o formato de metadados.

Quadro 4 – Metadados da canção Querelas do Brasil

Atributo	Conteúdo
Título	Querelas do Brasil
Ano da composição	1978
Compositor	Maurício Tapajós Aldir Blanc

Fonte: Elaborado pelo autor, a partir dos dados do encarte do disco

O documento utilizado para análise foi a gravação lançada em 1984, cujos dados de catalogação estão representados no Quadro 5:

Quadro 5 – Metadados da gravação de Querelas do Brasil, por Maurício Tapajós e Aldir Blanc

Atributo	Conteúdo
Disco	Aldir Blanc & Maurício Tapajós
Ano de lançamento	1984

Intérprete	Aldir Blanc Maurício Tapajós
Arranjador	Radamés Gnattali
Gravadora	SACI (Sociedade de Artistas e Compositores Independente)
Músicos	Aizik Meilach Geller: Violino Alceu de Almeida Reis: Violoncelo Aldir Blanc: Voz Antônio Adolfo: Piano Antônio Fidélis da Silva: Viola de Arco Arlindo Figueiredo Penteado: Viola de Arco Carlos Eduardo Hack: Violino Darcy Cruz: Trompete Ernani Bordinhão: Violino Frederick Stephany: Viola de Arco Genaldo Medeiros dos Santos: Saxofone Barítono Giancarlo Pareschi: Violino Hindemburgo Vitoriano Borges Pereira: Viola de Arco Jayme Araújo: Saxofone Alto João Daltro de Almeida: Violino José Alves da Silva: Violino José Dias de Lana: Violino Márcio Eymard Mallard: Violoncelo Maurício Tapajós: Voz Maurício Tapajós: Violão Sizão Machado: Baixo Elétrico Téo Lima: Bateria Walter Hack: Violino Zé Bodega: Saxofone Tenor

Fonte: Elaborado pelo autor, a partir dos dados do encarte do disco

A obtenção dos dados do documento gravação corresponde à leitura do nível imanente e instrumentaliza o postulado que afirma que a abordagem a domínios complexos deve partir do centro, para só então explorar as fronteiras a partir dos demais níveis de leitura, que no caso da semiótica musical, correspondem aos níveis poiético e estésico.

Outro elemento que pode ser acessado no nível imanente é o teor “em si” da gravação. Embora sempre que ouvido acione inevitavelmente o nível estésico daquele que a escuta, é possível considerar que há elementos estáveis, virtualmente disponíveis no registro da gravação. Entretanto, ao serem transpostos, seja em uma partitura musical, em uma tablatura de acordes ou textualmente pela letra da canção, esta transposição acarreta na perda substancial da capacidade da canção de produzir seus efeitos de sentido, uma vez que a linguagem sincrética é traduzida para outra linguagem, na qual não se pode acessar o enunciado de forma completa. Apesar desta ressalva, por limitações de suporte, será apresentada a seguir a transposição da letra da canção. Caso o leitor desta tese tenha a

possibilidade de ouvir a gravação, a mesma encontra-se disponível na web, no site Youtube, através do endereço: <https://www.youtube.com/watch?v=qVbqORhcb7I>⁹⁷

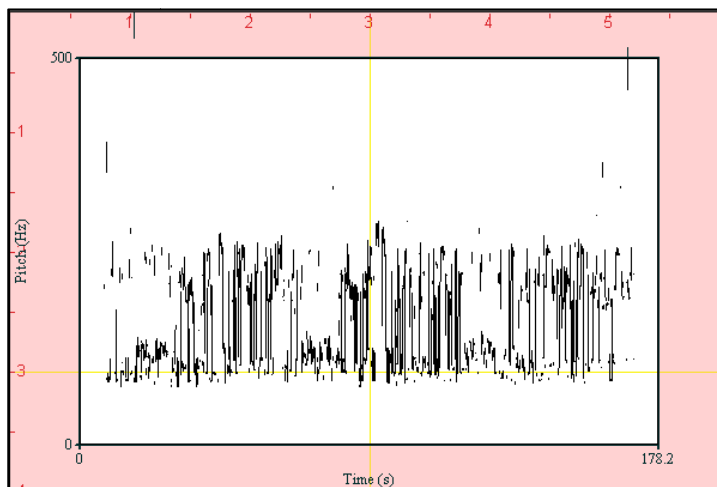
O Brazil não conhece o Brasil
 O Brasil nunca foi ao Brasil
 Tapir, jabuti, liana, alamandra, ali, alaúde
 Piau, ururau, aqui, ataúde
 Piá, carioca, porecramecrã
 Jobim akarore Jobim-açu
 Oh, oh, oh
 Pererê, camará, tororó, olererê
 Piriri, ratatá, karatê, olará
 O Brazil não merece o Brasil
 O Brazil tá matando o Brasil
 Jereba, saci, caandrades
 Cunhãs, ariranha, aranha
 Sertões, Guimarães, bachianas, águas
 E Marionaíma, ariraribóia,
 Na aura das mãos de Jobim-açu
 Oh, oh, oh
 Jererê, sarará, cururu, olerê
 Blablablá, bafafá, sururu, olará
 Do Brasil, SOS ao Brasil
 Do Brasil, SOS ao Brasil
 Do Brasil, SOS ao Brasil
 Tinhorão, urutu, sucuri
 Ujobim, sabiá, bem-te-vi
 Cabuçu, Cordovil, Cachambi, olerê
 Madureira, Olaria e Bangu, Olará
 Cascadura, Água Santa, Acari, Olerê
 Ipanema e Nova Iguaçu, Olará
 Do Brasil, SOS ao Brasil
 Do Brasil, SOS ao Brasil (TAPAJÓS; BLANC, 1988).

Da gravação é possível ainda extrair dados objetivos como modo, tonalidade, andamento e tessitura, e deles obter os coeficientes tensivos. Através da utilização de

⁹⁷ Acesso em: mai 2016.

ferramentas específicas de análise e síntese de arquivos sonoros, como o Praat⁹⁸, é possível realizar análises espectrais de aspectos fonológicos e fonéticos, conforme ilustra a Figura 25 que apresenta no eixo horizontal o tempo e no eixo vertical as alturas medidas em hertz.

Figura 22 – Análise da canção Querelas do Brasil, utilizando ferramenta Praat



Fonte: Adaptado da ferramenta Praat.

Recorrendo à tipologia proposta por Tatit, é possível identificar Querelas do Brasil como uma canção de perfil melódico preponderantemente tematizado, com grande repetição de motivos de curta duração, que permite a longa e reiterada enumeração de elementos expressos através da letra e que só produzem seus efeitos de sentido quando enunciados de forma progressiva e recorrente.

Considerando que a separação entre os níveis poiético e estésico não é absoluta, uma vez que a abordagem ao poiético passa obrigatoriamente pela (re)atualização do dispositivo infocomunicacional – o que ocorre no ato de se ouvir a canção –, as considerações tecidas a seguir sobre a análise do nível poiético são aproximações, deliberadamente simplificadas na tentativa de abordar intenções a partir dos vestígios materiais mais evidentes, como preconiza Jean Molino.

Como é frequente no cancioneiro popular brasileiro, a canção Querelas do Brasil dialoga com outras canções, autores e manifestações culturais. O título Querelas do Brasil

⁹⁸ Praat é um software gratuito para análise e síntese da fala a partir de arquivos digitais, desenvolvido pela Universidade de Amsterdã.

provavelmente remete à Aquarela do Brasil, canção de autoria de Ari Barroso⁹⁹ e grande sucesso desde seu lançamento em 1939. Contudo, a ressignificação provocada pela troca da palavra “aquarela” pela ambígua “querela” – que tanto pode significar discussão, debate; ou lamento, queixa; quanto denúncia; contenda – instaura uma diferenciação em relação à ufanista Aquarela de Ari Barroso, retrato das belezas e riquezas do Brasil.

A repetição tematizante da melodia permite aos autores de Querelas do Brasil explorar na letra diversas maneiras de dizer a mesma coisa, usando palavras em tupi, neologismos, gírias e nomes próprios constituindo uma intrincada rede de significados imanentes. Mais evidente, no entanto, é o mote trazido pela letra e reafirmado na gravação pela pronúncia que marca a diferença entre um Brasil real e um *Brazil* oficial, que não se reconhecem. Na última estrofe isso é evidenciado na enumeração de bairros e localidades do subúrbio carioca e da baixada fluminense, lugares preponderantemente pobres, sem apelo turístico ao contrário do famoso bairro de Ipanema que entra evidenciando a ruptura, ao mesmo tempo em que resume a concessão, a mistura e, possivelmente, a reconciliação entre o Brasil e o *Brazil*.

A escolha meticulosa das palavras, comum na obra de Aldir Blanc (mais um representante da mistura entre cancionista e poeta, Aldir é também escritor e médico formado em psiquiatria), constroi camadas de sentido no aparente *nonsense* discursivo. O *nonsense* é apenas aparente, pois ao citar elementos tipicamente brasileiros, porém desconhecidos da maioria do público, o cancionista vai comprovando a sua hipótese que o *Brazil* não conhece o Brasil. A parceria com Mauricio Tapajós (1943-1995), compositor, instrumentista, cantor e produtor musical brasileiro, declaradamente influenciado por Tom Jobim, traz elementos decisivos para a canção. Dentro da enumeração de elementos brasileiros, uma das poucas repetições, ainda que com variações, é a do termo “Jobim”, mencionado como “Jobim akarore”, “Jobim-açu” e “Ujobim”. Termos originais da língua tupi, *akarore* refere-se a índios gigantes; *açu* quer dizer grande; e o prefixo “U” quer dizer “o pai” (DALLE LUCCA, 2006). Dentro da consagração de elementos brasileiros que o Brasil oficial não reconhece ou não valoriza, os autores destacam Antonio Carlos Brasileiro de Almeida Jobim como ápice da brasilidade, da miscigenação entre as coisas do Brasil e a visão do *Brazil*. No plano da expressão melódica, esta deferência é enfatizada pela tonicidade, uma vez que o ponto mais extremo da tessitura, a nota mais aguda, ocorre justamente nas menções ao seu nome.

⁹⁹ Há controvérsias a respeito da coautoria de Aquarela do Brasil. O compositor Luís Peixoto (1889-1973), parceiro habitual de Ari Barroso, teria contribuído na composição, mas depois rejeitado a parceria por discordar de modificações realizadas por Ari. Para detalhes, ver: CABRAL, Sérgio. No tempo de Ari Barroso. Rio de Janeiro: Lumiar Editora, 1993.

Neste ponto da análise, os limites entre o estésico e o poiético não são mais definíveis, sendo relevante apontar que é justamente deste todo complexo que é possível desdobrar as relações que a canção popular oferece. Por exemplo, que o reconhecimento de Tom Jobim como “pai” ou “maioral” da música brasileira é recorrente. Este mesmo tipo de afirmação está presente em Paratodos, de Chico Buarque:

O meu pai era paulista
 Meu avô, pernambucano
 O meu bisavô, mineiro
 Meu tataravô, baiano
 Meu maestro soberano
 Foi Antonio Brasileiro (BUARQUE, 1993)

Deferência similar ocorre em Podres Poderes, de Caetano Veloso, onde Tom é evocado junto com Hermeto Pascoal e Milton Nascimento como possível salvador em relação ao obscurantismo e às trevas denunciadas pela canção:

Será que apenas os hermetismos pascoais
 Os toms, os miltons, seus sons e seus dons geniais
 Nos salvam, nos salvarão dessas trevas e nada mais? (VELOSO, 1984)

O princípio recursivo do método complexo evidencia que o desdobramento das relações é potencialmente infinito, uma vez que as relações são estabelecidas de dentro para fora (do centro complexo para as fronteiras que este mesmo desdobramento configura). É também no método complexo, mas a partir do princípio hologramático que reside a consistência de se considerar que, estando a origem no todo complexo, qualquer ponto desta rede permite a abordagem ao todo. Neste sentido, em vez de obrigar a explorar recursivamente todos os desdobramentos possíveis, um sistema de organização do conhecimento para domínios complexos pode exercer plenamente a sua natureza arbitrária e escolher quais relações irá desdobrar.

Em qualquer canção popular, as possibilidades são muitas. Em Querelas do Brasil, além de Tom Jobim, há referências explícitas ou indiretas a elementos da natureza, como animais (tapir, jabuti, ariranha, sabiá etc.) e a pessoas e obras da cultura nacional:

- Caandrades, escrito no plural, possivelmente referência aos irmãos Mário e Oswald de Andrade e, seguramente, a Carlos Drummond de Andrade;

- Marionaíma, amálgama de Mário de Andrade e seu personagem, o “herói sem nenhum caráter” Macunaíma, epítome do brasileiro;
- Sertões guimarães, referência a “Grande sertão: veredas” e a seu autor Guimarães Rosa;
- Bachianas águas, novamente a mistura entre o clássico e o popular, entre as Bachianas Brasileiras de Heitor Villa-Lobos e, presumivelmente, Águas de Março, de Tom Jobim.

A crítica ao *Brazil* que não conhece o Brasil é também apresentada com recurso a elementos típicos da cultura, como ocorre nos versos nos quais, após a reiteração do apelo “do Brasil SOS ao Brasil”, a menção a “tinhorão, urutu, sucuri”, possivelmente uma alusão ao controvertido crítico de música brasileira José Ramos Tinhorão juntamente com a menção a duas cobras peçonhentas; e também no neologismo “ariraribóia”, referência a Ari Barroso e ao índio Araribóia. Esta última relação de Querelas do Brasil com o autor da já mencionada Aquarela do Brasil, foi escolhida (sempre arbitrariamente) para ser explorada.

Ari de Resende Barroso (1903- 1964) foi pianista, compositor, apresentador, narrador de futebol e é considerado o criador do gênero “samba-exaltação”, durante a década de 1930, precisamente no período em que Getúlio Vargas empreendia seu projeto de criação de uma cultura nacional popular. Por este motivo, Ari gozava de grande prestígio político, embora no meio musical fosse por vezes considerado como um “vendido” ou “traidor” (CABRAL, 1993). A referência de Blanc e Tapajós ao índio Arariboia, portanto, não é fortuita, Arariboia foi chefe da tribo dos temiminós e os relatos históricos afirmam que teria ajudado os portugueses contra os tamoios e os franceses na luta pela conquista da baía de Guanabara em 1567, tendo sido recompensado com uma região de terra na entrada da baía, que atualmente corresponde à cidade de Niterói (KNAUSS, 2003).

Portanto, a querela apresentada por Blanc e Tapajós é um confronto com a visão oficial do “Brasil brasileiro” de Barroso. De forma análoga ao que foi mencionado para Querelas do Brasil, a seguir é apresentada a transcrição da letra da gravação original de Aquarela do Brasil. No entanto, para uma apreensão completa da análise, é sugerida a audição da gravação original, disponível no site YouTube ¹⁰⁰:

<https://www.youtube.com/watch?v=H-y8TS7jbpY>

¹⁰⁰ Acesso em: mai 2016.

Brasil
Meu Brasil brasileiro
Meu mulato inzoneiro
Vou cantar-te nos meus versos
Ô Brasil, samba que dá
Bamboleio que faz gingar
Ô Brasil, do meu amor
Terra de Nosso Senhor
Brasil, Brasil
Pra mim, pra mim

Ah, abre a cortina do passado
Tira a Mãe Preta, do serrado
Bota o Rei Congo, no congado
Brasil, Brasil
Pra mim, pra mim

Deixa, cantar de novo o trovador
A merencória luz da lua
Toda canção do meu amor
Quero ver a Sa Dona, caminhando
Pelos salões arrastando
O seu vestido rendado
Brasil, Brasil
Pra mim, pra mim

Brasil
Terra boa e gostosa
Da morena sestrosa
De olhar indiscreto
Ô Brasil, samba que dá
Bamboleio, que faz gingar
Ô Brasil, do meu amor
Terra de Nosso Senhor
Brasil, Brasil
Pra mim, pra mim

Oh, esse coqueiro que dá coco
Onde eu amarro a minha rede
Nas noites claras de luar
Brasil, Brasil

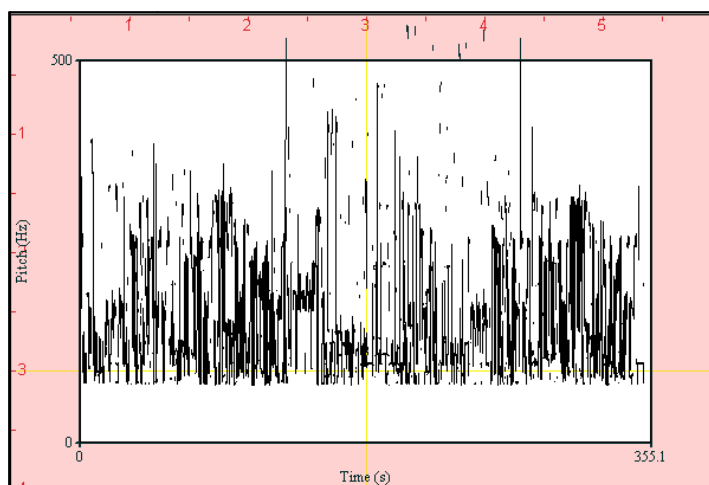
Pra mim, pra mim
 Ah, ouve essas fontes murmurantes
 Aonde eu mato a minha sede
 E onde a lua vem brincar
 Ah, este Brasil lindo e trigueiro
 É o meu Brasil, brasileiro
 Terra de samba e pandeiro
 Brasil, Brasil
 Pra mim, pra mim (BARROSO, 1939)

A gravação de Aquarela do Brasil em 1939 foi realizada pelo cantor Francisco Alves (1898-1952), com arranjo e regência de Radamés Gnattali, o mesmo arranjador de Querelas do Brasil em 1988. Esta coincidência, apesar de fortuita, não é uma exceção no âmbito da canção popular, onde a circulação dos músicos e suas redes de contatos e parcerias são reveladoras da mistura e das sobreposições. Sem desdobrar mais uma relação, mas no intuito de ilustrar a afirmação anterior, pode-se mencionar que o mesmo Radamés Gnattali é compositor de diversas obras de música clássica (série Brasileiras; Concertos para piano e orquestra; Concerto para Violoncelo e orquestra; e Concerto para harmônica de boca, por exemplo) e é o pianista e personagem enaltecido na canção Quatro Loucos Num Samba, de Ciro Monteiro (1913-1973):

Oba!
 O samba vai começar
 Quatro loucos fazendo miséria
 Na bossa legal sem parar
 O Radamés dá um breque de piano para começar
 [...] (MONTEIRO, 1962)

Retornando a “Aquarela do Brasil”, a análise espectral da gravação de Francisco Alves revela um perfil melódico distinto da gravação de Querelas do Brasil, confirme ilustra a Figura 23.

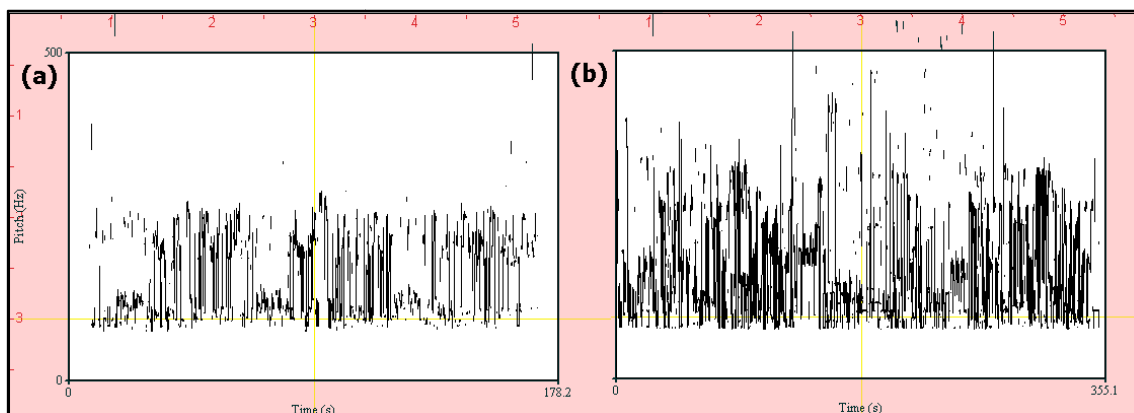
Figura 23 – Análise da canção Aquarela do Brasil, utilizando ferramenta Praat



Fonte: Adaptado da ferramenta Praat.

Sendo igualmente caracterizada pela recorrência melódica típica das canções tematizadas, esta gravação de Aquarela do Brasil, no entanto, apresenta uma tessitura expandida em relação a Querelas do Brasil, indicando elementos de passionalização, condizente com os arroubos da interpretação de Francisco Alves que alonga praticamente todas vogais das sílabas tônicas, possivelmente com a intenção de conferir dramaticidade à interpretação. A Figura 24 apresenta as duas imagens lado a lado, para facilitar a comparação visual das duas análises.

Figura 24 – Comparação das análises das canções Querelas do Brasil (a) e Aquarela do Brasil (b)



Fonte: Adaptado da ferramenta Praat.

Do ponto de vista do conteúdo, a relação entre as duas canções é notória, uma vez que abordam o mesmo tema, contudo, a partir de pontos de vista distintos. Querelas do Brasil

questiona a imagem que se faz do Brasil e evoca os personagens e elementos considerados típicos da cultura com a intenção de promover um resgate do Brasil em pleno período de abertura política da Ditadura Militar (a composição é de 1978). Em *Aquarela do Brasil*, por outro lado, como o próprio título sugere, é apresentado um retrato estático da terra de “Nosso Senhor”, naturalmente rica, bela, harmônica, pacífica e miscigenada, como afirmam os versos que definem o país como “mulato inzoneiro”, “lindo e trigueiro”, adequado aos anseios políticos do regime ditatorial do Estado Novo.

Estas análises, que desde o início não se pretendiam exaustivas, reafirmam a complexidade da canção popular e a necessidade de soluções específicas para sua abordagem através de sistemas de organização do conhecimento. Através da descrição em linguagem em OWL e da definição de propriedades fuzzy, tornou-se possível recuperar e tratar este tipo de objeto de maneira mais adequada.

6.3 CONSTRUÇÃO DO SISTEMA E RESULTADOS ALCANÇADOS

A construção do sistema teve início pela transformação dos metadados de catalogação apresentados nos Quadros 4 e 5 em um esquema escrito em RDF, utilizando a definição em triplas, ou seja, considerando a descrição semântica das relações entre atributos e valores, conforme ilustra o trecho de Código fonte 5, que apresenta a descrição de uma canção e seu compositor.

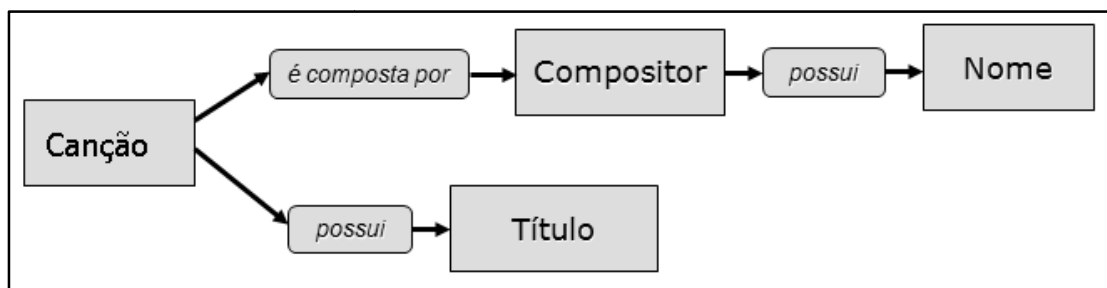
Código fonte 5 – Definição de canção e compositor, expressa em RDF

```
<rdfs:Class rdf:ID="cancao"/>
<rdfs:Class rdf:ID="compositor"/>
<rdfs:ObjectProperty rdf:ID="eCompostaPor">
  <rdfs:domain rdf:resource="#cancao"/>
  <rdfs:range rdf:resource="#compositor"/>
</rdfs:ObjectProperty>
<rdfs:DatatypeProperty rdf:ID="possuiTitulo">
  <rdfs:domain rdf:resource="#cancao"/>
  <rdfs:range rdf:resource="string"/>
</rdfs:DatatypeProperty>
<rdfs:DatatypeProperty rdf:ID="possuiNome">
  <rdfs:domain rdf:resource="#compositor"/>
  <rdfs:range rdf:resource="string"/>
</rdfs:DatatypeProperty>
```

Fonte: Elaborado pelo autor.

A Figura 25 apresenta uma ilustração gráfica da descrição de compositor e título de uma canção expressas em RDF.

Figura 25 – Ilustração visual da descrição de canção e compositor, expressa em RDF



Fonte: Elaborado pelo autor.

Conforme detalhado anteriormente, a descrição em RDF baseia-se em triplas, que permite que os objetos possam ser sujeitos em outros predicados, criando uma rede de significados com semântica explícita. No exemplo apresentado na Figura 25, *Compositor* é o sujeito na relação *Compositor possui Nome* ao mesmo tempo em que é objeto na relação *Canção é composta por Compositor*.

A separação entre a estrutura e os dados é promovida pelo framework RDF, uma vez sujeitos e objetos são materializados a partir da criação de instâncias, por exemplo, Aldir Blanc, Maurício Tapajós e Ari Barroso são instâncias de *Compositor*; Querelas do Brasil e Aquarela do Brasil são instâncias de *Canção*.

O trecho de Código fonte 6 apresenta a criação de instâncias para a definição da canção Querelas do Brasil e seus compositores.

Código fonte 6 – Definição das instâncias da canção Querelas do Brasil em RDF

```

<compositor rdf:ID="http://www.exemplo.com.br/cancao/aldir-blanc">
  <rdfs:label>Aldir Blanc</rdfs:label>
  <possuiNome rdf:datatype="string">Aldir Blanc</possuiNome>
</compositor>

<compositor rdf:ID="http://www.exemplo.com.br/cancao/mauricio-tapajos">
  <rdfs:label>Maurício Tapajós</rdfs:label>
  <possuiNome rdf:datatype="string">Maurício Tapajós</possuiNome>
</compositor>

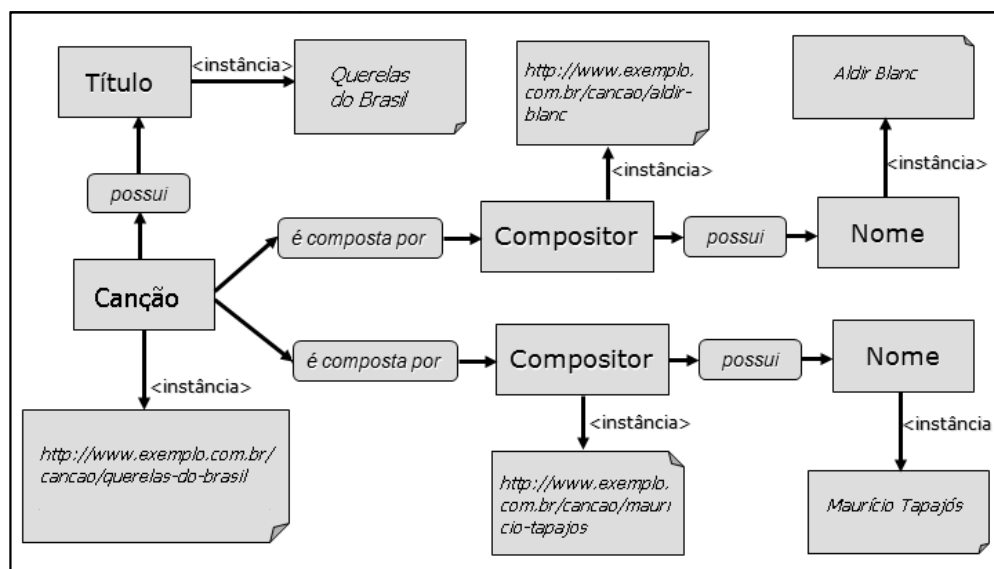
<cancao rdf:ID="http://www.exemplo.com.br/cancao/querelas-do-brasil">
  <rdfs:label>Querelas do Brasil</rdfs:label>
  <possuiTitulo rdf:datatype="string">Querelas do
Brasil</possuiTitulo>
  <eCompostaPor rdf:resource="http://www.exemplo.com.br/cancao/
mauricio-tapajos" />
  <eCompostaPor rdf:resource="http://www.exemplo.com.br/cancao/aldir-
blanc" />
</cancao>

```

Fonte: Elaborado pelo autor.

A Figura 26 apresenta o mesmo exemplo visualmente, destacando-se em *itálico* as instâncias.

Figura 26 – Ilustração da descrição em RDF com instâncias, para a canção Querelas do Brasil e seus compositores



Fonte: Elaborado pelo autor.

Através da utilização do RDF é possível indicar ao computador a semântica dos relacionamentos entre os recursos, porém não se alcança a realização do raciocínio automático. Para permitir sua realização, foi necessário estender o modelo utilizando os recursos da linguagem OWL. Sem esta extensão, seria obrigatório, por exemplo, declarar cada possível papel de uma pessoa explicitamente:

- Ari Barroso é um compositor;
- Ari Barroso compôs Aquarela do Brasil;
- Radamés Gnattali é compositor;
- Radamés Gnattali compôs Concerto para Violoncelo e orquestra;
- Radmés Gnattali é arranjador;
- Radamés Gnattali arranjou Aquarela do Brasil;
- Radamés Gnattali arranjou Querelas do Brasil etc.

Através da definição de regras em OWL é possível declarar cada pessoa uma única vez e inferir automaticamente o seu papel de acordo com as propriedades às quais está relacionado. Por exemplo, uma pessoa que tem alguma canção compostaPor ela, é inferida automaticamente como um Compositor. Uma pessoa que tem alguma gravação arranjadaPor ela, é automaticamente inferida como um Arranjador, conforme demonstrado no Código fonte 7.

Código fonte 7 – Descrição dos dados de canção, gravação, composição e intérprete, expressos em OWL

```

<owl:Class rdf:about="http://www.exemplo.com.br/cancao/cancao"/>
<owl:Class rdf:about="http://www.exemplo.com.br/cancao/pessoa"/>
<owl:Class rdf:about="http://www.exemplo.com.br/cancao/gravacao"/>

<owl:Class rdf:about="http://www.exemplo.com.br/cancao/compositor">
  <owl:equivalentClass>
    <owl:Restriction>
      <owl:onProperty rdf:resource="http://www.exemplo.com.br/
cancao/compostaPor"/>
      <owl:someValuesFrom rdf:resource="http://www.exemplo.
com.br/cancao/pessoa"/>
    </owl:Restriction>
  </owl:equivalentClass>
  <rdfs:subClassOf rdf:resource="http://www.exemplo.com.br/
cancao/pessoa"/>
</owl:Class>

<owl:Class rdf:about="http://www.exemplo.com.br/cancao/arranjador">
  <owl:equivalentClass>
    <owl:Restriction>
      <owl:onProperty rdf:resource="http://www.exemplo.com.br/
cancao/arranjadaPor"/>
      <owl:someValuesFrom rdf:resource="http://www.exemplo.com.
br/cancao/pessoa"/>
    </owl:Restriction>
  </owl:equivalentClass>
  <rdfs:subClassOf rdf:resource="http://www.exemplo.com.br/cancao/
pessoa"/>
</owl:Class>

<owl:ObjectProperty
rdf:about="http://www.exemplo.com.br/cancao/gravada">
  <rdfs:domain rdf:resource="http://www.exemplo.com.br/cancao/
cancao"/>
  <rdfs:range rdf:resource="http://www.exemplo.com.br/cancao/
gravacao"/>
</owl:ObjectProperty>

owl:ObjectProperty rdf:about="http://www.exemplo.com.br/cancao/
compostaPor">
  <rdfs:domain rdf:resource="http://www.exemplo.com.br/cancao/
cancao"/>
  <rdfs:range rdf:resource="http://www.exemplo.com.br/cancao/
pessoa"/>
</owl:ObjectProperty>

<owl:ObjectProperty rdf:about="http://www.exemplo.com.br/cancao/
arranjadaPor">
  <rdfs:domain rdf:resource="http://www.exemplo.com.br/cancao/
gravacao"/>
  <rdfs:range rdf:resource="http://www.exemplo.com.br/cancao/
pessoa"/>
</owl:ObjectProperty>

```

Fonte: Elaborado pelo autor.

Ainda em relação aos dados de catalogação, foram definidas regras que permitem informar as relações entre *Músico* e *Instrumento* (*tocarInstrumento*) e entre *tocarInstrumento* e *Gravação* (*tocouEm*). Os músicos que tocam juntos em uma mesma gravação estabelecem uma propriedade simétrica *tocouJuntoCom* e, para cada tipo de instrumento foi declarada uma função que permite obter o designativo de quem o toca, por exemplo, uma pessoa que tocou flauta em alguma gravação é um flautista.

Outra regra identificada foi a que determina a parceria entre compositores, explicitando que, quando uma mesma música é *compostaPor* mais de um compositor, eles estabelecem uma relação simétrica de *parceria*, conforme ilustra o Código fonte 8:

Código fonte 8 – Declaração da regra parceria, em OWL

```
<owl:SymmetricObjectProperty rdf:about="http://www.exemplo.com.br/cancao/parceria">
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="#compostaPor"/>
      <owl:allValuesFrom>
        <owl:Class>
          <owl:unionOf rdf:parseType="Collection">
            <owl:Class rdf:about="#Compositor"/>
          </owl:unionOf>
        </owl:Class>
      </owl:allValuesFrom>
    </owl:Restriction>
  </rdfs:subClassOf>
</owl:SymmetricObjectProperty>
```

Fonte: Elaborado pelo autor.

Conforme mencionado anteriormente, a utilização da OWL permite a integração de dados a partir da explicitação de relação entre conteúdos de outras fontes, como o projeto DBPedia, por exemplo. O Código fonte 9 apresenta o trecho de código utilizado para realizar esta ligação entre os dados de Radamés Gnattali e Aldir Blanc com suas respectivas descrições na DBPedia. O exemplo a seguir ilustra também a já discutida característica de sistema aberto alcançada com a OWL: o artista Maurício Tapajós não possui recurso associado na DBPedia, podendo ser igualmente descrito como uma pessoa sem que isto signifique sua inexistência ou incompletude, como ocorreria em linguagens computacionais tradicionais, nas quais a omissão significa, necessariamente, inexistência.

Código fonte 9 – Definição de dados ligados, promovendo a integração com recursos do projeto DBPedia

```
<owl:NamedIndividual rdf:about="#Radamés_Gnattali">
  <rdf:type rdf:resource="http://www.exemplo.com.br/cancao/pessoa"/>
  <rdfs:label>Radamés Gnattali</rdfs:label>
  <owl:sameAs rdf:resource="http://dbpedia.org/resource/
Radam%C3%A9s_Gnattali"/>
</owl:NamedIndividual>

<owl:NamedIndividual rdf:about="#Aldir_Blanc">
  <rdf:type rdf:resource="http://www.exemplo.com.br/cancao/pessoa"/>
  <rdfs:label>Aldir Blanc</rdfs:label>
  <owl:sameAs rdf:resource="http://dbpedia.org/resource/
Aldir_Blanc"/>
</owl:NamedIndividual>

<owl:NamedIndividual rdf:about="#Maurício_Tapajós">
  <rdf:type rdf:resource="http://www.exemplo.com.br/cancao/pessoa"/>
  <rdfs:label>Maurício Tapajós</rdfs:label>
</owl:NamedIndividual>
```

Fonte: Elaborado pelo autor.

Apesar dos avanços alcançados através desta forma de descrição no que diz respeito às possibilidades de descrição de um dispositivo infocomunicacional e sua rede de relações recursiva e potencialmente infinita, processável de forma automática através das inferências e aplicação de regras lógicas, as relações ainda se resumem a declaração ou não das regras. A definição de propriedades fuzzy é necessária para permitir expressar graus de pertencimento das instâncias às regras.

A Figura 27 ilustra a criação de uma propriedade fuzzy na ferramenta Protégé ¹⁰¹ para expressar os perfis melódicos (passionalização, tematização e figurativização). Os valores de intervalo atribuídos para representar os graus de pertinência a cada perfil foram escolhidos arbitrariamente (0, 0.5 e 1.0; 0.5, 1.0 e 1.5; 1.5, 2.0 e 2.5, respectivamente) de forma a representar níveis de gradação e possibilidade de sobreposição entre os perfis.

¹⁰¹ Software gratuito para edição de “ontologias computacionais”, desenvolvido pela Universidade de Stanford. A extensão que permite definir dados fuzzy foi aplicada a partir das definições de STRACCIA (2013).

Figura 27 – Definição de propriedades fuzzy para expressar os graus de pertinência aos perfis melódicos

The image shows the Protégé interface for defining a fuzzy property. It is divided into two steps:

- Step 1: Select the datatype to annotate**
 - Enter a datatype name:
 - + Add new datatype
 - Selected datatype: **tematica**
- Step 2: Choose type and parameters**
 - Type: **triangular** (dropdown menu)
 - Parameters:
 - A:
 - B:
 - C:

To the right of the parameters, there is a graph illustrating a triangular fuzzy set. The x-axis is labeled 'x' and has points 0, a, b, and c. The y-axis has a value of 1. The graph shows a triangle with vertices at (a, 0), (b, 1), and (c, 0). A dashed vertical line connects the peak at (b, 1) to the x-axis at point b.

Fonte: Adaptado da ferramenta Protégé.

Com este tipo de propriedade é possível exprimir em que grau uma determinada canção possui a característica temática, podendo, simultaneamente possuir algum outro grau de pertencimento às propriedades de passionalização e figurativização. Estes valores de pertencimento são obtidos a partir do cálculo dos coeficientes tensivos e permitem descrever em um formato computacionalmente processável os resultados das análises realizadas com a ferramenta Praat e ilustrados através das Figuras 22 e 23.

O Código fonte 10 apresenta a criação da propriedade fuzzy relativa ao perfil melódico tematização. As demais propriedades são análogas, variando apenas os valores limites da função *fuzzyfication*.

Código fonte 10 – Definição de propriedades fuzzy para expressar os graus de pertinência aos perfis melódicos, em OWL

```

<owl:DatatypeProperty
  rdf:about="http://www.exemplo.com.br/cancao/tematizacao">
<rdfs:label>tematizacao</rdfs:label>
<rdfs:domain>
  <owl:Restriction>
    <owl:onProperty
      rdf:resource="http://www.exemplo.com.br/cancao/tematizacao"/>
    <owl:minQualifiedCardinality rdf:datatype="http://www.w3.org/
2001/XMLSchema#nonNegativeInteger">1</owl:minQualifiedCardinality>
    <owl:onDataRange rdf:resource="http://www.exemplo.com.br/
cancao/tematica"/>
  </owl:Restriction>
</rdfs:domain>
</owl:DatatypeProperty>

<rdfs:Datatype rdf:about="http://www.exemplo.com.br/cancao/tematica">
<owl:equivalentClass>
  <rdfs:Datatype>
    <owl:intersectionOf rdf:parseType="Collection">
      <rdfs:Datatype>
        <owl:onDatatype rdf:resource="http://www.w3.org/
2001/XMLSchema#double"/>
        <owl:withRestrictions rdf:parseType="Collection">
          <rdf:Description>
            <xsd:minInclusive rdf:datatype="http://www.w3.
org/2001/XMLSchema#double">0.5</xsd:minInclusive>
          </rdf:Description>
        </owl:withRestrictions>
      </rdfs:Datatype>
      <rdfs:Datatype>
        <owl:onDatatype rdf:resource="http://www.w3.org/2001/
/XMLSchema#double"/>
        <owl:withRestrictions rdf:parseType="Collection">
          <rdf:Description>
            <xsd:maxInclusive rdf:datatype="http://www.w3.
org/2001/XMLSchema#double">1.5</xsd:maxInclusive>
          </rdf:Description>
        </owl:withRestrictions>
      </rdfs:Datatype>
    </owl:intersectionOf>
  </rdfs:Datatype>
</owl:equivalentClass>

<fuzzyLabel>
  <fuzzyOwl2 fuzzyType='datatype'>
    <Datatype type='triangular'; a='1.0' b='1' c='1.5' />
  </fuzzyOwl2>
</fuzzyLabel>
</rdfs:Datatype>

```

Fonte: Elaborado pelo autor.

A obtenção dos dados de catalogação e a análise de Querelas do Brasil discutida na seção anterior, resultou na identificação do conjunto de classes apresentado no Quadro 6.

Quadro 6 – Classes identificadas a partir da abordagem empírica

Classe	Descrição
Canção	Classe que se refere aos dados de uma canção popular brasileira. Exemplo: Aquarela do Brasil
Pessoa	Classe que se refere a uma pessoa. Exemplo: Tom Jobim
Instrumento	Classe que se refere a um instrumento musical. Exemplo: Flauta
Perfil Melódico	Classe que se refere a um Perfil Melódico. Possui três valores pré-definidos: Tematização; Passionalização e Figurativização
Gravadora	Classe que se refere a uma Gravadora, responsável por uma gravação. Exemplo: RCA Victor
Regime Político	Classe que representa um determinado regime político. Exemplo: Estado Novo.
Tema	Classe que representa um tema mencionado pela letra. Exemplo: Brasil
Gravação	Classe que representa uma gravação, a interpretação de uma canção. Exemplo: a gravação de Francisco Alves para Aquarela do Brasil

Fonte: Elaborado pelo autor.

As propriedades identificadas e descritas no sistema construídos estão listadas no Quadro 7.

Quadro 7 – Propriedades identificadas a partir da abordagem empírica

Propriedade	Fuzzy	Descrição
possuiTitulo	não	Relaciona um título a uma canção
possuiNome	não	Relaciona um nome a uma classe qualquer (canção, pessoa, instrumento, perfil melódico, gravadora, regime político ou tema), exceto gravação
compostaEm	não	Relaciona uma data de composição a uma canção
gravadaEm	não	Relaciona uma data de gravação a uma gravação
pelagravadora	não	Relaciona uma gravadora a uma gravação
letra	não	Relaciona uma letra a uma canção
éParceiroDe	não	Relaciona um compositor a um compositor
éArranjadorDe	não	Relaciona um arranjador a uma gravação
éCompositorDe	não	Relaciona um compositor a uma gravação
éIntérpreteDe	não	Relaciona um intérprete a uma gravação
tocouEm	não	Relaciona um músico a uma gravação
tocouCom	não	Relaciona um músico a um músico
gravadaDurantePeriodoPolitico	não	Relaciona uma gravação a um período político
possuiPerfilMelodicoTematização	sim	Relaciona um grau de pertencimento a um perfil melódico de tematização
possuiPerfilMelodicoPassionalização	sim	Relaciona um grau de pertencimento a um perfil melódico de passionalização
possuiPerfilMelodicoFigurativização	sim	Relaciona um grau de pertencimento a um perfil melódico de figurativização
possuiTemática	sim	Relaciona um tema a uma canção, definindo o grau de pertinência
fazMencaoA	sim	Relaciona uma temática abordada a uma canção, definindo o grau de pertinência

fazMencaoElogiosaA	sim	Relaciona uma temática abordada a uma canção, a partir de um grau de elogio
fazMencaoCriticaA	sim	Relaciona uma temática abordada a uma canção, a partir de um grau de crítica

Fonte: Elaborado pelo autor.

E, finalmente, as regras identificadas e descritas no sistema estão listadas no Quadro 8.

Quadro 8 – Regras identificadas a partir da abordagem empírica

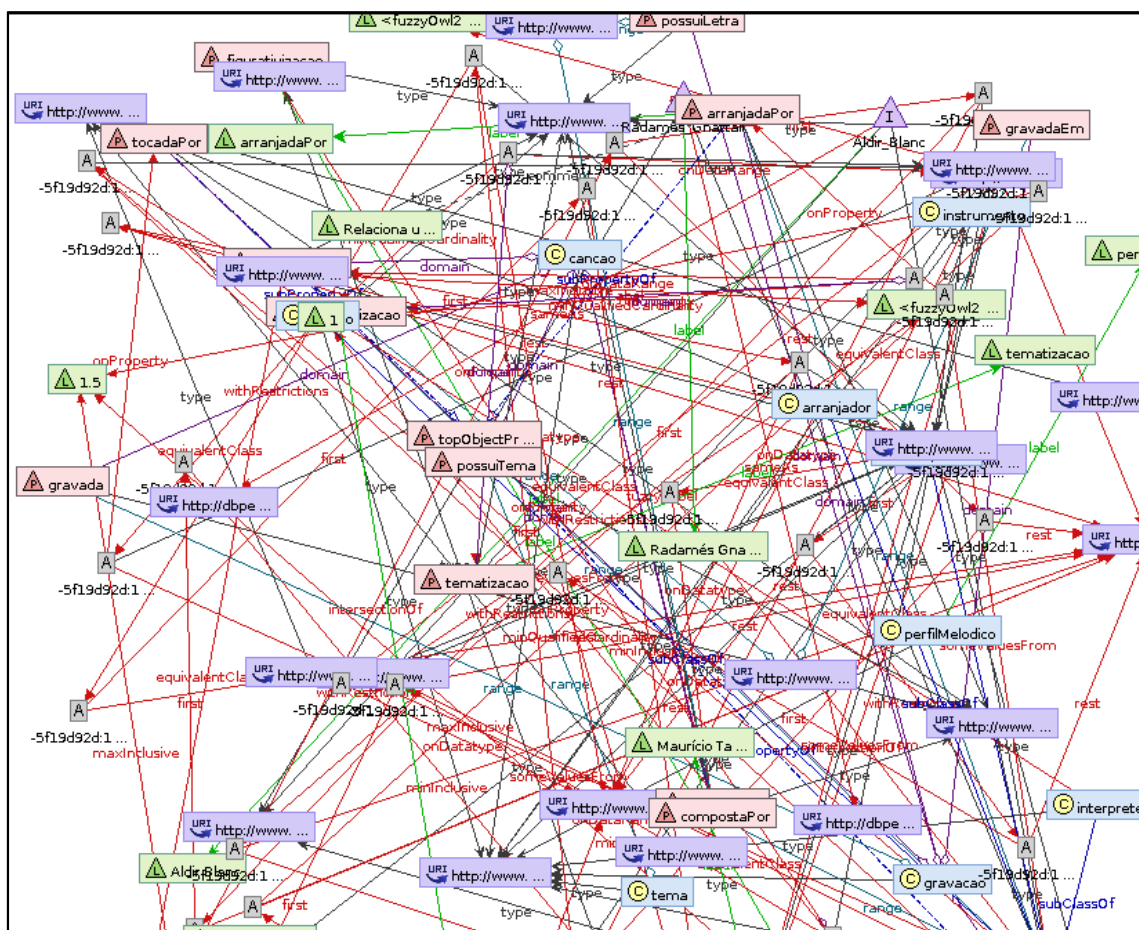
Regra	Descrição
Intérprete	Regra que define um intérprete como uma pessoa que éInterpreteDe uma gravação
Compositor	Regra que define um compositor como uma pessoa que éCompositorDe uma canção
Arranjador	Regra que define um arranjador como uma pessoa que éArranjdorDe uma gravação
Músico	Regra que define um intérprete como uma pessoa que tocouEm uma gravação
Parceria	Regra que define que duas pessoas são compositoras da mesma canção
Período Histórico	Regra que define que uma canção que foi gravada durante determinado regime, pertence a um período histórico
Temática Relacionada	Regra que define a temática relacionada entre as canções a partir de suas menções
TocouJuntoCom	Regra que define que dois músicos que tocaram juntos em uma determinada gravação
GravadaPor	Regra que define que uma canção foi gravada

Fonte: Elaborado pelo autor.

A criação de instâncias correspondentes aos dados obtidos na análise da canção Querelas do Brasil em suas classes, propriedades e regras, resulta em uma estrutura rizomática que é ilustrada pela Figura 28, através da visualização gerada com a utilização da ferramenta *Gravity*¹⁰².

¹⁰² Ferramenta gratuita, desenvolvida pelo Salzburg Research's Knowledge Information Systems Group.

Figura 28 – Visão rizomática do sistema de organização do conhecimento construído

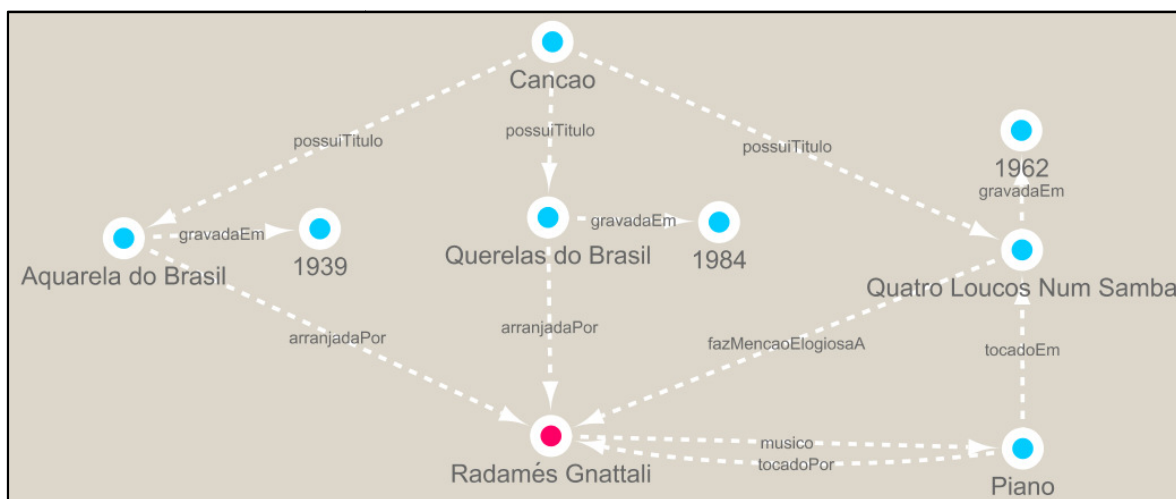


Fonte: Adaptado da ferramenta Gravity.

Esta figura sintetiza a visão do todo complexo, a partir do qual é possível selecionar dimensões de análise para serem desdobradas e assim conduzir a recuperação da informação. É, portanto, a recuperação quem conduz as estratégias de representação e da própria estrutura, uma vez que a ativação de regras dentro do modelo OWL será realizada conforme a recuperação desejada.

A Figura 29 ilustra a recuperação realizada a partir da seleção de o nó correspondente a Radamés Gnattali.

Figura 29 – Visão rizomática do sistema a partir da recuperação de dados relativos a Radamés Gnattali



Fonte: Elaborado pelo autor.

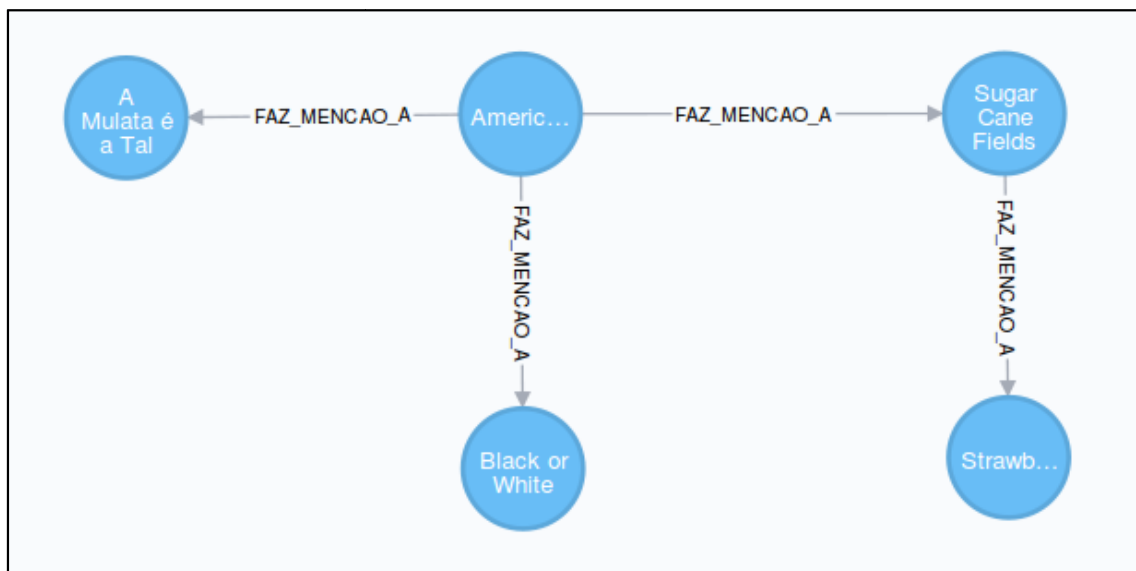
Na recuperação exemplificada pela Figura 29, são apresentadas as três gravações relacionadas a Radamés Gnattali, e seu papel em cada uma delas (arranjador em Aquarela do Brasil e em Querelas do Brasil; e como homenageado e como pianista em Quatro Loucos Num Samba).

Utilizando uma interface de navegação na web, o sistema de organização do conhecimento desenvolvido permite aos seus utilizadores dirigir a tarefa de recuperação. Esta recuperação pode se basear tanto na navegação a partir das classes, correspondendo em certa medida à visão arbórea dos sistemas tradicionais que se baseiam em relações de subordinação; quanto a partir dos relacionamentos e regras, correspondendo, neste caso, à visão rizomática do conhecimento.

Um protótipo de interface de navegação é apresentado a seguir e aplicado ao exemplo retomado do capítulo 3 – a canção “Americanos” de Caetano Veloso e suas relações.

Na Figura 30 é apresentada a navegação a partir da classe Canção, explicitando os desdobramentos de outras canções a partir de “Americanos”.

Figura 30 – Interface web. Navegação pela classe Canção a partir de “Americanos”

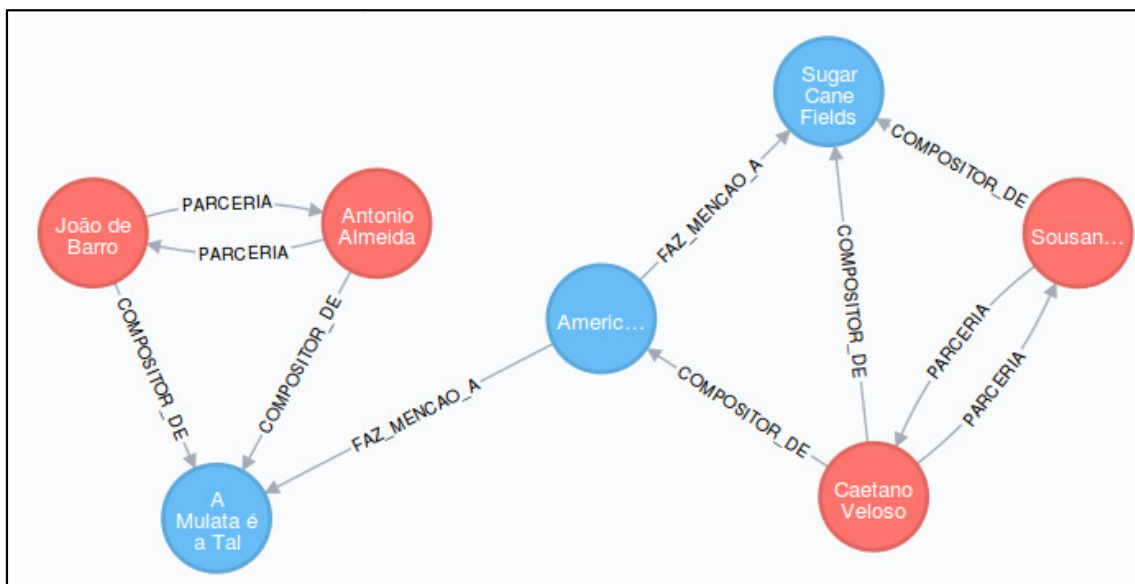


Fonte: Elaborado pelo autor.

Nesta navegação, foram encontradas as canções “Americanos” e aquelas mencionadas por ela: “A Mulata é a Tal”, “Black or White” e “Sugar Cane Fields Forever”. Em um segundo nível de subordinação, foi identificada a canção “Strawberry Fields Forever”, mencionada a partir de “Sugar Canes Fields Forever”.

A visão rizomática pode ser alcançada se, em lugar das estruturas de classes, a recuperação for realizada a partir das relações e regras, como exemplificado na Figura 31, na qual a navegação é realizada a partir da relação “Compositor de” que recupera as pessoas que compuseram canções e que, eventualmente, estabelecem entre si uma relação de parceria, identificada automaticamente a partir da aplicação da regra correspondente.

Figura 31 – Interface web. Navegação pela relação “Compositor de”, explicitando a regra “Parceria”

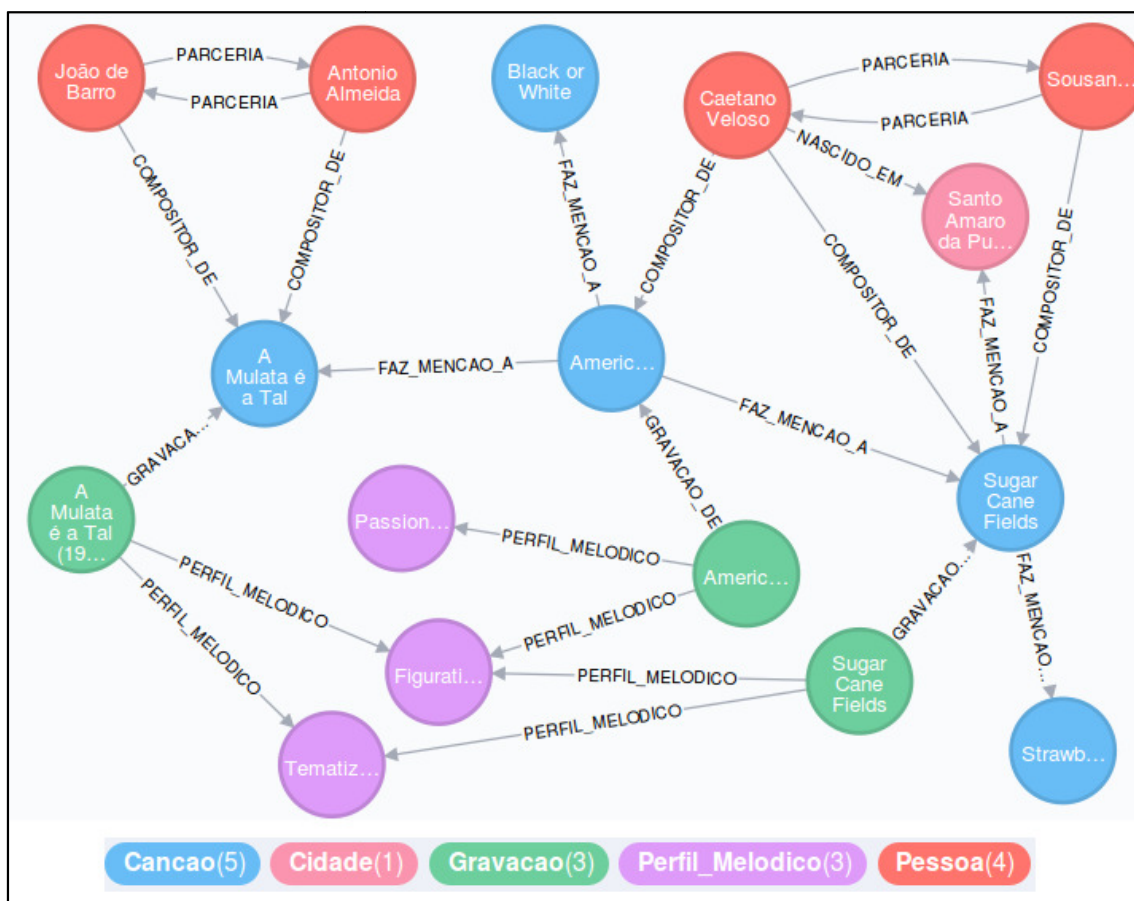


Fonte: Elaborado pelo autor.

Nesta navegação, foram recuperadas apenas as canções “Americanos”, “A Mulata é a Tal” e “Sugar Cane Fields Forever”. Este resultado é ilustrativo da visão de sistemas abertos anteriormente mencionada, uma vez que as propriedades “Compositor De” não foram descritas para as canções “Strawberry Fields Forever” e “Black or White”.

No protótipo desenvolvido, o desdobramento é feito através de hiperlinks que permitem selecionar qualquer nó da rede e expandi-lo. Na Figura 32 é exemplificado o desdobramento das canções e de suas gravações.

Figura 32 – Interface web. Navegação pelas relações desdobradas a partir de “Americanos”

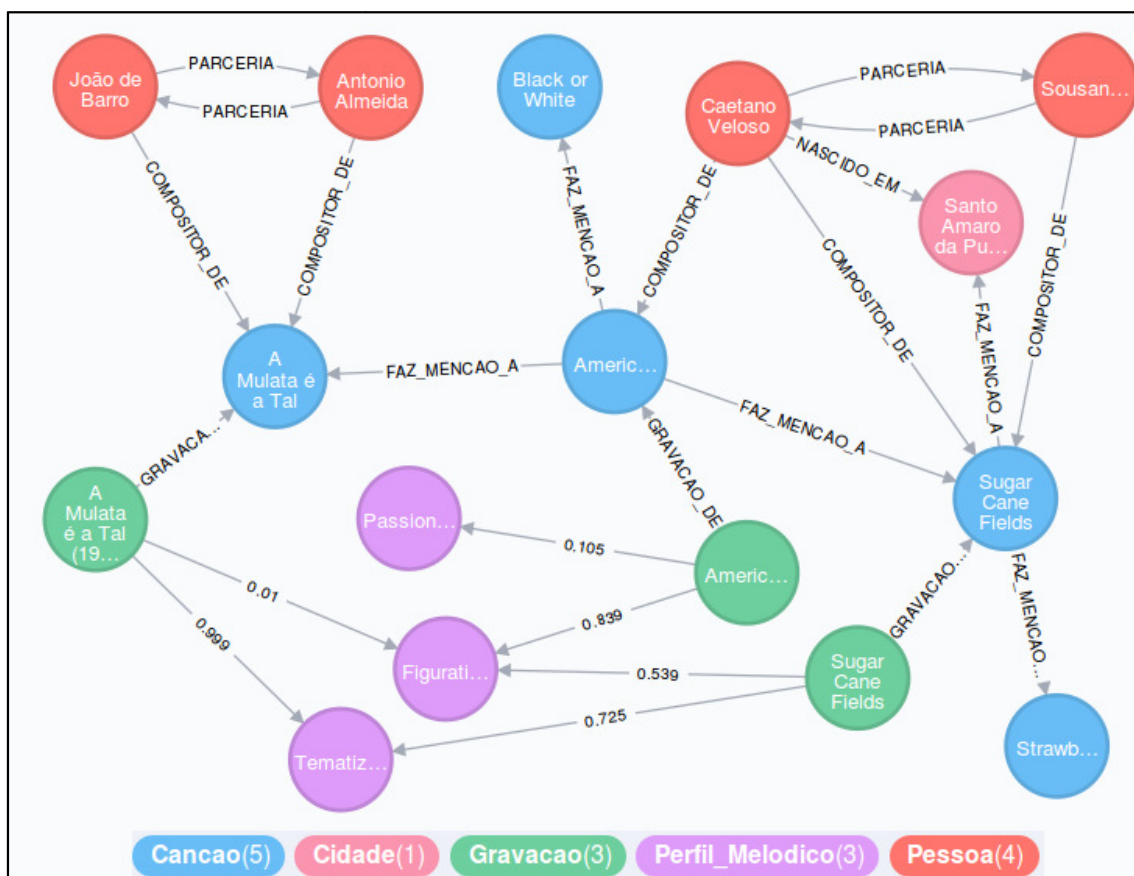


Fonte: Elaborado pelo autor.

A descrição semântica das relações alcançadas através do uso da linguagem OWL permite explicitar as propriedades, como, por exemplo: o ano e o intérprete de cada gravação; ou o ano de composição de cada canção, de forma que o computador torna-se capaz de “entender” e processar estas relações de forma automática.

As propriedades mensuráveis em graus, como é o caso dos perfis melódicos, foram descritas utilizando propriedades fuzzy e podem ser recuperadas a partir de seus valores expressos em graus de pertencimento. Na Figura 33 é apresentado o exemplo da medição dos coeficientes tensivos das gravações de “A Mulata é a Tal”, “Americanos” e “Sugar Cane Fields Forever”.

Figura 33 – Interface web. Navegação explicitando os graus de pertinência dos perfis melódicos das gravações



Fonte: Elaborado pelo autor.

Embora limitado pelo suporte quando expresso de forma estática nesta tese, o sistema desenvolvido revelou a possibilidade de uma navegação simultaneamente multidimensional e dinâmica, evidenciando que, através da primazia da recuperação, os resultados obtidos através dos dispositivos infocomunicacionais, improváveis nos sistemas de organização do conhecimento tradicionais, podem ser multiplicados ao limite da criatividade dos utilizadores de um sistema desenvolvido considerando a abordagem a domínios complexos.

7 CONCLUSÃO

A conclusão desta pesquisa é articulada em duas frentes complementares. Na primeira, o intuito é verificar criticamente a pertinência dos resultados obtidos em relação aos objetivos enunciados. A segunda frente diz respeito à identificação de novas questões de pesquisa e possíveis desdobramentos decorrentes do trabalho realizado, sendo apontados através de propostas de trabalhos futuros.

O levantamento teórico realizado neste trabalho buscou identificar as bases adequadas para elaboração de um sistema de organização do conhecimento capaz de abordar domínios complexos. A apresentação de uma revisão histórica teve por objetivo evidenciar a anterioridade da necessidade de elaboração deste tipo de sistema de organização, refletida pelas suas formulações ao longo dos séculos, como nas obras de conde de Buffon, passando por Ranganathan e chegando a pensadores contemporâneos da Organização do Conhecimento.

Seguindo a tradição continental europeia, esta pesquisa considerou a concepção semiótica do conhecimento – notadamente a representação simbólica do conhecimento – como forma de abordagem ao real, a partir da obra de Ernst Cassirer. O primado da percepção e sua conciliação com a inteligência como forma de organizar o mundo é por si incompatível com o modelo científico tradicional – simplificador e dicotômico –, sendo necessária uma teoria capaz de considerar a complexidade não como um problema, mas como possibilidade.

A teoria de abordagem ao pensamento complexo desenvolvida por Edgar Morin foi utilizada em conjunto com construtos teóricos de outros pensadores da “escola francesa”. Reconhecendo a insuficiência da noção de “conceito” formulado enquanto unidade de conhecimento, uma vez que o conhecimento passa a ser considerado como complexo (simultaneamente produto e produtor da sua teia de relações), foi tomada a noção de dispositivo segundo a concepção original de Michel Foucault e suas releituras por Giorgio Agamben e Gilles Deleuze. No tocante à Organização do Conhecimento, foi considerada a formulação apresentada por Regina Marteleto e Viviane Couzinet, de dispositivo infocomunicacional como o elo sensível constituído a partir das relações entre elementos que que concorrem em um mesmo espaço tensivo, no qual produzem efeitos de sentido.

A obtenção da síntese, ponto culminante de qualquer processo dialético, uma vez que representa o momento de equilíbrio das oscilações que a produz, é alcançada pela restituição da centralidade da noção de documento. No entanto, considera-se um documento “vivo”,

simultaneamente como suporte que permite a circulação e a abordagem ao dispositivo com suas relações imanentes, reveladoras do caráter simbólico do conhecimento.

Do ponto de vista metodológico, a proposta de um sistema de organização do conhecimento a partir deste referencial teórico foi compatibilizada a partir da adoção da noção de rizoma no lugar de estruturas hierárquicas e do consequente deslocamento da compreensão de fronteira de lugar de distinção e separação para um lugar de integração, de sobreposição e de multiplicidades, conforme proposto pelos filósofos Gilles Deleuze e Félix Guattari. Esta operação se baseou no Método do Pensamento Complexo, proposto por Edgar Morin, que fornece os princípios operativos para abordagem ao complexo dentro de um modelo baseado em estruturas. No entanto, as estruturas são consideradas não como oposições dicotômicas e sim como construções baseadas em oscilações e em gradações. Esta formulação foi alcançada pelo recurso ao esquematismo tensivo, a partir da formulação original de Viggo Brøndal e dos desenvolvimentos contemporâneos liderados por Claude Zilberbeg.

Uma vez formulada a concepção de sistemas de organização do conhecimento não dicotômicos, com fronteiras sobrepostas, fundados na noção de dispositivos infocomunicacionais e cujas relações são mensuráveis em graus, buscou-se, ainda no que concerne ao aspecto metodológico, ferramentas e ambientes adequados para sua implementação. Foram discutidos o avanço da computação – no que diz respeito ao aumento da capacidade de processamento de dados como fator de transformação das possibilidades para a criação de sistemas de organização do conhecimento. A capacidade do computador de “simular” e processar a não dicotomia a partir da multiplicação potencialmente infinita de subdivisões entre os valores zero e um (a base do bit, unidade indivisível de processamento computacional), permitiu sistematizar e realizar de maneira automática processamentos de lógica fuzzy, um tipo de lógica não clássica, surgida na segunda metade do século XX, a partir dos trabalhos de Lotfi Zadeh.

Esta expansão da capacidade de processamento, que poderia se resumir a um “mero ganho de escala” das técnicas de automação, foi acompanhado pelo surgimento da web, primeiramente como ambiente de compartilhamento de documentos interligados – na forma de páginas e hiperlinks –, e atualmente como plataforma de produção de conteúdo em escala global. A chamada “explosão de dados” causada pela web teve como uma de suas muitas consequências a descentralização dos repositórios de armazenamento – os dados estão distribuídos, poder-se-ia dizer “espalhados” pela rede – e trazem a necessidade de sistemas que possam interligar estes dados e fornecer respostas precisas dentro de um volume a cada dia maior de dados. Estes sistemas vem sendo pensados e construídos a partir da retomada de

duas noções fulcrais para a Organização do Conhecimento. A primeira diz respeito à centralidade da recuperação em lugar da representação, conforme preconizado por Jesse Shera a partir da década de 1950. A segunda diz respeito à necessidade que esta recuperação seja realizada de forma automática pelo computador a partir de regras e princípios operativos formalizados, conforme vaticinado por outro teórico da Organização do Conhecimento, Brian Vickery, ainda na década de 1980. A aplicação destas noções está em curso no âmbito do projeto denominado web semântica, conduzido por Tim Berners-Lee e o consórcio W3C, e que objetiva permitir que o computador conduza de forma automática a tarefa de recuperação, operando sobre representações e estruturas heterogêneas. Neste intuito, foram apresentadas as linguagens e tecnologias da web semântica que embasaram a construção empírica de um sistema de organização do conhecimento aplicado a um domínio complexo: o domínio das canções populares brasileiras.

Preliminarmente, foram discutidos os aspectos teóricos que configuram a canção popular como um domínio autônomo do conhecimento, especialmente em relação a outros tipos de expressões artísticas musicais. Explicitou-se que esta autonomia se deve ao fato da canção popular possuir uma linguagem própria, uma linguagem sincrética – na medida em que mistura a entoação da voz que fala, a melodia da voz que canta e o texto do que é falado/cantado. Foram discutidos os antecedentes sociais e políticos que transformaram a canção popular brasileira em uma forma de expressão artística de amplo espectro na cultura brasileira, das raízes miscigenadas à sua apropriação enquanto projeto político.

Os aspectos metodológicos relacionados à canção popular incluíram a investigação sobre o uso de metadados e de métodos colaborativos e também a análise de suas abordagens na web e na web semântica. Foram investigados projetos e iniciativas de organização do conhecimento em música – especificamente em canção popular –, e discutidas suas potencialidades e limitações. Por fim, foram apresentados estudos relativos à abordagem semiótica e tensiva da canção popular que buscam estabelecer ferramentas de abordagem específicas para este domínio do conhecimento.

A construção empírica de um sistema de organização do conhecimento para canções populares baseada no referencial teórico e metodológico apresentado mostrou-se realizável e promissora. Ao alcançar os objetivos estabelecidos pela pesquisa, os resultados obtidos confirmaram as hipóteses de viabilidade de abordar um domínio complexo considerando a multidimensionalidade do conhecimento a partir de fronteiras sobrepostas e relações de pertencimento expressas em graus. O deslocamento da base da Organização do Conhecimento da noção de conceito para a noção de dispositivo infocomunicacional mostrou-se adequada

para lidar com a complexidade ao permitir o acesso às relações e não apenas aos atributos – as condições necessárias e suficientes dos modelos tradicionais. Estas relações, expressas através de propriedades fuzzy, permitiram operacionalizar o esquematismo tensivo e abordar o conhecimento complexo. Os novos recursos computacionais, como as linguagens RDF e OWL da web semântica permitiram desenvolver um sistema real, integrando os dados da pesquisa a dados já existentes na web, realizando processamentos automáticos e confirmando a centralidade da atividade de recuperação para este tipo de sistema.

Finalmente, a pesquisa desenvolvida buscou atender a uma demanda de investigações de soluções para sistemas de organização do conhecimento que considerem os desenvolvimentos teóricos e tecnológicos recentes. Neste sentido, é possível considerar os objetivos plenamente alcançados e os resultados obtidos como sendo satisfatórios.

A pesquisa realizada e os resultados alcançados abrem possibilidades de continuidades e desdobramentos em diferentes áreas do conhecimento. De forma coerente com a concepção “complexa” da Organização do Conhecimento adotada nesta tese, serão enfatizadas as possibilidades que a colocam em diálogo diferentes abordagens e disciplinas.

- **Ligação de dados:** seguindo os princípios da web semântica, um trabalho futuro mais imediato decorrente desta tese é publicar na web os dados produzidos por esta pesquisa, ligando-os a recursos digitais já existentes, como, por exemplo, os do portal “Musica Brasilis”, e, posteriormente, a filiação desta iniciativa ao grupo de trabalho *Linked Data* da W3C, com o intuito de proporcionar a interoperabilidade semântica dos recursos sobre música brasileira na web.
- **Primitivos emocionais:** além das descrições e propriedades apresentadas no capítulo anterior, é possível prever a expansão do modelo para incorporar primitivos emocionais, como as noções de “humor” decorrentes dos coeficientes tensivos definidos por Carmo Jr e discutidos no capítulo 5.
- **Temporalidade:** outra possível expansão refere-se à incorporação de aspectos de temporalidade. Esta dimensão temporal atuaria sobre as propriedades possibilitando, por exemplo, recuperar artistas que foram parceiros até uma determinada época e depois brigaram; gravadoras que existiram até uma data e

depois foram extintas ou incorporadas por outras; artistas que trocaram de nome, como Jorge Ben que virou Jorge Ben Jor em 1989, por exemplo.

- **Interfaces de uso:** um trabalho futuro diz respeito a realizar estudos de usabilidade, investigando modelos e formatos de construção de interfaces de uso específicas para o sistema proposto. Estas interfaces devem prever tanto a visão do classificacionista quanto a do classificador ¹⁰³, considerando ainda as especificidades do ambiente web e dos aspectos colaborativos de construção do conhecimento.
- **Níveis integrativos:** a Teoria dos Níveis Integrativos, desenvolvida paralelamente por Nicolai Hartmann (1882-1950) e James Feibleman (1904-1987) ¹⁰⁴, pressupõe a organização hierárquica do conhecimento a partir de fenômenos dispostos em níveis sucessivos de especificidade. A despeito da aparente incompatibilidade com a visão rizomática adotada nesta pesquisa, um trabalho futuro poderia se beneficiar de alguns dos princípios estabelecidos pelos autores, especialmente da noção de “dependência existencial”, como forma de abordagem ao complexo (DE SANTIS; GNOLI, 2016).
- **Outras abordagens teóricas:** Outro trabalho futuro diz respeito à investigação de aproximações e diferenças entre a teoria proposta e outras abordagens teóricas e filosóficas e os aportes teóricos e metodológicos advindos destas investigações. Um exemplo possível é a aproximação à “Fenomenologia da Percepção” de Maurice Merleau-Ponty (1908-1961) e à noção de experiência, já presente nos fundamentos do esquematismo tensivo. Outra possibilidade seria a aproximação à “Lógica do Sentido” de Gilles Deleuze, que propõe ultrapassar as conformações empíricas e fenomenológicas do “sentido como acontecimento”, propondo um “terceiro estado do sentido” – simultaneamente transcendental e complexo, indiferente a todos os opostos (DELEUZE, 1969). Este caminho, possivelmente, permitiria fundamentar uma abordagem

¹⁰³ Utiliza-se aqui a terminologia proposta por Ranganathan: “classificador” é a pessoa que utiliza um sistema de classificação; enquanto “classificacionista” é a pessoa que propõe e elabora um sistema de classificação.

¹⁰⁴ Não há indícios que algum dos autores tenha tomado conhecimento dos desenvolvimentos realizados pelo outro.

“praxiológica do conhecimento”, apontada pelo presente autor e por Rosali Fernandez de Souza em um trabalho anteriormente realizado a partir da constatação bourdieusiana da falibilidade das abordagens de cunho subjetivista e objetivista da ciência (DE SANTIS; SOUZA, no prelo).

- **Análise Formal de Conceitos:** a possibilidade de colocar esta pesquisa em diálogo com as teorias cognitivistas muito provavelmente não prescindirá de explorar os trabalhos basilares de Eleanor Rosch e de George Lakoff. Destes autores, pioneiros dos estudos da Psicologia Cognitiva e da Linguística Cognitiva, respectivamente, advem as bases para teorias matemáticas do conhecimento como a “Análise Formal de Conceitos” ¹⁰⁵, proposta por Rudolf Wille na década de 1980 e que prevê a especificação de conceitos através de formalismos matemáticos, permitindo a construção de estruturas do conhecimento no formato de mapas conceituais cujo intuito é representar a gênese e permitir a manipulação dos conceitos em uma abordagem pretensamente próxima àquela realizada pela mente humana.
- **Espaços conceituais:** a proposta de representação geométrica do pensamento proposta pela teoria dos espaços conceituais ¹⁰⁶ (GÄRDENFORS, 2004) visam superar a noção materialista de “linguagem mental” (como em Ockham, por exemplo) bem como o formalismo matemático da Análise Formal de Conceitos. Para seu propositor, Peter Gärdenfors, os espaços conceituais são conjuntos de dimensões, nos quais cada dimensão é uma ordem de percepções reais possíveis, representáveis geometricamente através de formas (em lugar de axiomas lógicos) e apreensíveis pelas relações de proximidade ou distância entre os conceitos. O diálogo entre este tipo de abordagem cognitivista e a apresentada neste trabalho, entretanto, deve ser cuidadosa no sentido de não forçar aproximações intimamente incompatíveis sem o devido elo metodológico. Por este motivo, a Organização do Conhecimento não pode prescindir do recurso a formulações específicas para lidar com aspectos

¹⁰⁵ Do original, em inglês, *Formal Concept Analysis*.

¹⁰⁶ Do original, em inglês, *Conceptual Spaces – The Geometry of Thought*

sociais, tais como as mediações infocomunicacionais e a metodologia de Análise de Redes Sociais, mencionadas a seguir.

- **Mediação infocomunicacional:** um desdobramento desejável deste trabalho remete a estudos que visem incorporar ao sistema proposto a noção de mediação infocomunicacional, definida por Yves Jeanneret como “um terceiro simbólico”, e que permitiria expandir o modelo apresentado para abordar um conjunto de valores e de práticas compartilhadas que transcende o cotidiano das trocas simbólicas (JEANNERET, 2008). Estes estudos permitirão aproximar esta pesquisa de estudos da Comunicação que foram deliberadamente negligenciados em decorrência da escolha de privilegiar a especificidade do funcionamento das formas simbólicas.
- **Análise de Redes Sociais:** estabelecer o diálogo entre o sistema elaborado nesta tese e os estudos e ferramentas da Análise de Redes Sociais permitiriam aprofundar a abordagem das relações entre as pessoas ou organizações (agentes) ao mesmo tempo em que forneceria subsídios teóricos a respeito da estruturação destas redes (métricas de conexão, distribuição e segmentação, por exemplo) (DEGENNE; FORSÉ, 1994). A explicitação semântica das relações através do RDF e da OWL combinada com medidas oferecidas por programas computacionais de análise como, por exemplo, UCInet, permitiriam mensurar mais precisamente as relações de densidade ou de centralidade de grau de um artista, por exemplo.

REFERÊNCIAS

REFERÊNCIAS BIBLIOGRÁFICAS

ADOMAVICIUS, Gediminas; TUZHILIN, Alexander. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *Knowledge and Data Engineering, IEEE Transactions on*, v. 17, n. 6, p. 734–749, 2005.

AGAMBEN, Giorgio. *Qu'est-ce qu'un dispositif*. Paris: Rivages, 2007.

AGHAEI, Sareh; NEMATBAKHSH, Mohammad Ali; FARSANI, Hadi Khosravi. Evolution of the world wide web: from WEB 1.0 TO WEB 4.0. *International Journal of Web & Semantic Technology*, v. 3, n. 1, p. 1, 2012.

ANDRADE, Mário. *Ensaio sobre a música brasileira*. São Paulo: Martins, 1962.

AUSTIN, Derek. Two steps forward. In: PALMER, BERNARD I. *Itself an education*. London: The Libray Association, 1971. p. 69–110.

BACHELARD, Gaston. *Le Nouvel esprit scientifique*. Paris: PUF, 1949.

BARRETO, Aldo de Albuquerque. Mitos e lendas da informação: o texto, o hipertexto e o conhecimento. *Pesquisa Brasileira em Ciência da Informação e Biblioteconomia*, v. 2, n. 2, 2008.

BĚLOHLÁVEK, Radim *et al.* Concepts and fuzzy sets: misunderstandings, misconceptions, and oversights. *International Journal of Approximate Reasoning*, v. 51, n. 1, p. 23–34, 2009.

BERNERS-LEE, Tim; FISCHETTI, Mark. *Weaving the web: the original design and ultimate destiny of the World Wide Web by its inventor*. Nova Iorque: Harper, 1999.

BERNERS-LEE, Tim; HENDLER, James; LASSILA, Ora. The semantic web. *Scientific american*, v. 284, n. 5, p. 28–37, 2001.

BOURDIEU, Pierre. *Ce que parler veut dire: l'économie des échanges linguistiques*. Paris: Fayard, 1982.

BOURDIEU, Pierre. *O poder simbólico*. Rio de Janeiro: Bertrand Brasil, 1989.

BRØNDAL, Viggo. *Essais de linguistique générale*. København: Ejnar Munksgaard, 1943.

BUCKLAND, Michael. *Briet, 1894-1989, "Madame Documentation"*. Buckland website. Disponível em: <<http://people.ischool.berkeley.edu/~buckland/briet.html>>. Acesso em mar 2016. [s.l.][s.n.], 2007

BURKE, Peter; BRIGGS, Asa. *Uma história social da mídia: de Gutenberg à Internet*. Rio de Janeiro: Jorge Zahar, 2004.

CABRAL, Sérgio. *No tempo de Ari Barroso*. Rio de Janeiro: Lumiar Editora, 1993.

CAMPOS, Augusto; CAMPOS, Haroldo. *Revisão de Sousândrade*. Rio de Janeiro: Editora Nova Fronteira, 1982.

CARMO JR., José Roberto. Coeficientes melódicos da emoção na melodia tonal. In: SIMPÓSIO DE COGNIÇÃO E ARTES MUSICAIS, 10, 2014. Campinas.

CARMO JR., José Roberto. Eu te amo – canção de Tom Jobim e Chico Buarque. *CASA: Cadernos de Semiótica Aplicada*, v. 3, n. 1, 2005.

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

CASSIRER, Ernst. *A filosofia das formas simbólicas*. São Paulo: Martins Fontes, 2001b. v. 3.

CASSIRER, Ernst. *Linguagem e mito*. São Paulo: Perspectiva, 1985.

CHARTIER, Roger. *A história cultural: entre práticas e representações*. Lisboa: Difusão, 1988.

CHUNG, Ji-Moon. A study of future internet applications based on semantic web technology configuration model. *Indian Journal of Science and Technology*, v. 8, n. 20, 2015.

CICERO, Antonio. *Finalidades sem fim: ensaios sobre poesia e arte*. Rio de Janeiro: Editora Companhia das Letras, 2005.

COELHO, Márcio. A paixão e a passionalização em “Saudosa maloca”. *CASA: Cadernos de Semiótica Aplicada*, v. 1, n. 2, 2003.

COUZINET, Viviane. Complexidade e documento: a hibridação das mediações nas áreas em ruptura. *Revista Eletrônica de Comunicação, Informação & Inovação em Saúde*, v. 3, n. 3, p. 10–16, 2009.

COUZINET, Viviane. De la communication scientifique à la médiation spécialisée: communication des savoirs et formes d’hybridations. In: PAPY, Fabrice. *Problématiques émergentes dans les sciences de l’information*. Paris: Hermès, Lavoisier, 2008. p. 57–85.

DAHLBERG, Ingetraut. Brief communication: what is knowledge organization? *Knowledge Organization*, v. 41, n. 1, 2014.

DAHLBERG, Ingetraut. Teoria do conceito. *Ciência da informação*, v. 7, n. 2, p. 101–107, 1978.

DALLE LUCCA, Jussara. Querelas do Brasil. In: SIMPÓSIO DE MÚSICA DA FAP, 2, 2006, Curitiba.

DAM CHRISTENSEN, Hans. The framing of scientific domains: about UNISIST, domain analysis and art history. *Journal of Documentation*, v. 70, n. 2, p. 261–281, 2014.

DAY, Michael. The long-term preservation of Web content. *Web Archiving*. Springer, 2006. p. 177–199.

DE SANTIS, Rodrigo. *Aquela música: navegação interativa por canções populares brasileiras*. Rio de Janeiro: Funarte, 2010. (Funarte - Fundação Nacional das Artes).

DE SANTIS, Rodrigo. Knowledge Organization and the conceptual basis for building classification systems for complex documents: an application on the Brazilian popular song domain. *Advanced Information Systems Engineering Workshops*, pp. 349-360. Springer Berlin Heidelberg, 2013.

DE SANTIS, Rodrigo; GNOLI, Claudio. Expressing dependence relationships in the Integrative Levels Classification using OWL. In: *International ISKO Conference*, 14, 2016, Rio de Janeiro. No prelo.

DE SANTIS, Rodrigo; LANZELOTTE, Rosana Saldanha da Gama. Difusão de partituras através da web: o desafio de tornar acessíveis os repertórios brasileiros. In: GOLDBERG, Guilherme; SANTOS, Luís Otávio; VALVERDE, Rodolfo. *Intertextualidades: fronteiras entre o sacro e o profano na música do Brasil colonial e imperial*. Juiz de Fora: Museu de Arte Murilo Mendes, 2014. p. 288–323.

DE SANTIS, Rodrigo; SOUZA, Rosali Fernandez de. A indefinição é o regime: a classificação de canções populares em gêneros musicais. CONGRESO DE LA ASOCIACIÓN LATINOAMERICANA DE SOCIOLOGÍA (ALAS), 29, 2013, Santiago, Chile. Disponível em: <<http://actacientifica.servicioit.cl/>>. Acesso: nov 2015.

DE SANTIS, Rodrigo; SOUZA, Rosali Fernandez de. Uma proposta de diálogo entre o conhecimento praxiológico de Bourdieu e a Organização do Conhecimento. In: MARTELETO, Regina M.; PIMENTA, Ricardo M. *Pierre Bourdieu e a produção social da cultura, da informação e do conhecimento*. [S.l.]. No prelo.

DEGENNE, Alain; FORSÉ, Michel. *Les réseaux sociaux: une approche structurale en sociologie*. Paris: Armand Collin, 1994.

DELALANDE, François. Music analysis and reception behaviours: Sommeil by Pierre Henry. *Journal of new music research*, v. 27, n. 1–2, p. 13–66, 1998.

DELEUZE, Gilles. *Différence et répétition*. Paris: PUF, 1968.

DELEUZE, Gilles. *Logique du sens*. Paris: Minuit, 1969.

DELEUZE, Gilles. O atual e o virtual. Allez, Éric. *Deleuze filosofia virtual*. Rio de Janeiro: 34, 1996.

DELEUZE, Gilles. Qu'est-ce qu'un dispositif? *Michel Foucault philosophe: rencontre internationale*. Paris, 1988. v. 9. p. 316–325.

DELEUZE, Gilles; GUATTARI, Félix. Introdução: rizoma. *Mil platôs: capitalismo e esquizofrenia*, v.1. São Paulo: Editora 34, 1995.

DELEUZE, Gilles; GUATTARI, Félix. *Qu'est-ce que c'est la philosophie*. Paris: Minuit, 1980.

DIETRICH, Peter. *Araçá Azul: uma análise semiótica*. 2003. Dissertação (Mestrado). Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo. São Paulo, 2003.

DIETRICH, Peter. Discurso musical e discurso de produção musical. *CASA: Cadernos de Semiótica Aplicada*, v. 5, n. 2, 2007.

DIMAGGIO, Paul. Classification in art. *American Sociological Review*, p. 440–455, 1987.

DIMAGGIO, Paul *et al.* Social implications of the Internet. *Annual Review of Sociology*, p. 307–336, 2001.

DODEBEI, Vera Lucia Doyle Louzada de Mattos. Continuidades e rupturas em representação do conhecimento. In: DODEBEI, Vera; GUIMARÃES, José Augusto Chaves. *Complexidade e organização do conhecimento: desafios de nosso século*. Estudos Avançados em Organização do Conhecimento. Rio de Janeiro: ISKO Brasil, 2013. v. 2, p. 79–85.

DOWNIE, J. Stephen; BYRD, Donald; CRAWFORD, Tim. *Ten years of ISMIR*: reflections on challenges and opportunities Kobe, ISMIR. 2009, p. 13–18.

ECO, Umberto. *Da árvore ao labirinto*: estudos históricos sobre o signo e a interpretação. Rio de Janeiro: Record, 2013.

FEDELI, Gian Carlo. Metaphors of order and disorder: from the tree to the labyrinth and beyond. *Knowledge Organization*, v. 40, n. 6, 2013.

FIORIN, José Luiz. *As astúcias da enunciação*: as categorias de pessoa, espaço e tempo. São Paulo: Editora Ática, 1996.

FONTANILLE, Jacques; ZILBERBERG, Claude. *Tensão e significação*. São Paulo: Discurso Editorial: Humanitas/FFLCH/USP, 2001.

FOSKETT, Douglas J. Beyond reductionism. *Journal of Librarianship and Information Science*, v. 2, n. 2, p. 139–143, 1970.

FOUCAULT, Michel. *L'Archéologie du savoir*. Paris: Gallimard, 1969.

FOUCAULT, Michel. Le jeu de Michel Foucault. [Entretien avec D. Colas, A. Grosrichard, G. Le Gaufey, J. Livi, G. Miller, J. Miller, J.-A. Miller, C. Millot, G. Wajeman]. *Bulletin Périodique du Champ Freudien*, n. 10, p. 62–93, 1977.

FOUCAULT, Michel. *L'Ordre du discours*. Paris: Gallimard, 1971.

FOUCAULT, Michel. *Surveiller et punir*. Paris: Gallimard, 1975.

FRANCESCHI, Humberto Moraes. *A Casa Edison e seu tempo*. São Paulo: Sarapuí, 2002.

FREYRE, Gilberto. *Casa-grande & senzala*: formação da família brasileira sob o regime da economia patriarcal. Rio de Janeiro: José Olympio, 1933.

GÄRDENFORS, Peter. *Conceptual spaces*: the geometry of thought. Massachusetts: MIT press, 2004.

GNOLI, Claudio; SZOSTAK, Rick. The León Manifesto. *Knowledge Organization*, v. 34, n. 1, p. 6–8, 2007.

GOMES, Hagar Espanha. Tendências da pesquisa em Organização do Conhecimento. *Pesquisa Brasileira em Ciência da Informação*, v. 2, n. 1, p. 60–88, 2009.

GONZÁLEZ DE GÓMEZ, Maria Nélide. A representação do conhecimento e o conhecimento da representação: algumas questões epistemológicas. *Ciência da Informação*, v. 22, n. 3, p. 217–222, 1993.

GONZÁLEZ DE GÓMEZ, Maria Nélide. O caráter seletivo das ações de informação. *Informare - Cadernos Programa de Pós Graduação em Ciência da Informação*, v. 5, n. 2, p. 7–31, 1999.

GUAN, Chun; YUEN, Kevin Kam Fung; COENEN, Frans. Towards an intuitionistic fuzzy agglomerative hierarchical clustering algorithm for music recommendation in folksonomy. *IEEE*, p. 2039–2042, 2015.

HOLT, Jen. Nécrologie de Viggo Brøndal. *Acta Linguistica*, v. 3, n. 1, 1943.

HJØRLAND, Birger; ALBRECHTSEN, Hanne. Toward a New Horizon in Information Science: Domain-Analysis. *Journal of the American Society for Information Science*, n. 46, v.6, p. 400-425, 1995.

HJØRLAND, Birger. Domain analysis in information science. Eleven approaches - traditional as well as innovative. *Journal of Documentation*, n.58, v.4, p.422-462, 2002.

JACOB, Elin K. Classification and categorization: a difference that makes a difference. *Library Trends*, v. 52, n. 3, p. 515–540, 2004.

JEANNERET Yves. *Penser la trivialité*, tome 1, la vie triviale des êtres culturels, Paris: Hermès, 2008.

KEITH, Douglas K. Displaying implicit associations among items in loosely-structured data sets. In: *Google Patents*, 2003.

KLIR, George J.; ST. CLAIR, Ute H.; YUAN, Bo. *Fuzzy set theory: foundations and applications*. New Jersey: Prentice Hall, 1997.

KNAUSS, Paulo. Herói da Cidade – imagem indígena e mitologia política. In: KNAUSS, Paulo. *Sorriso da cidade: imagens urbanas e história política de Niterói*. Niterói: Niterói Livros, 2003.

LANGRIDGE, Derek. *Classificação - abordagem para estudantes de biblioteconomia*. Rio de Janeiro: Interciência, 1977.

LANZELOTTE, Rosana S. G. *Música secreta: minha viagem ao Brasil, 1816-1821*. Rio de Janeiro: Editora Arte Ensaio, 2009.

LARSEN, Svend Erik. Introduction. Viggo Brøndal (1887-1942), linguiste, philosophe, sémioticien. *Langages* (A l'occasion d'un centenaire: l'Actualité de Brøndal). v. 22, n. 86, p. 7–12, 1987.

MAMDANI, Ebrahim H.; ASSILIAN, Sedrak. An experiment in linguistic synthesis with a fuzzy logic controller. *International journal of man-machine studies*, v. 7, n. 1, p. 1–13, 1975.

MARIZ, Vasco. *A canção brasileira: erudita, folclórica, popular*. Brasília: Livraria Editora Cátedra em convênio com o Instituto Nacional do Livro, Ministério da Educação e Cultura, 1977.

MARTELETO, Regina Maria; COUZINET, Viviane. Mediações e dispositivos de informação e comunicação na apropriação de conhecimentos: elementos conceituais e empíricos a partir de olhares intercruzados. *Revista Eletrônica de Comunicação, Informação & Inovação em Saúde*, v. 7, n. 2, p. 1–16, 2013.

MAZZOCCHI, Fulvio. Images of thought and their relation to classification: the tree and the net. *Knowledge Organization*, v. 40, n. 6, p. 366–74, 2013.

MAZZOCCHI, Fulvio. Western science and traditional knowledge. *EMBO reports*, v. 7, n. 5, p. 463–466, 2006.

MEYRIAT, Jean. Document, documentation, documentologie. l'écrit et le document. *Schéma et schématisation*, v. 14, p. 51–63, 1981.

MICHAELIS INGLÊS-PORTUGUÊS. In: MICHAELIS. *Michaelis: dicionário inglês-português, português-inglês*. 70. ed. São Paulo: Melhoramentos, 2002. p. 796.

MOLINO, Jean. Fait musical et semiologie de la musique. *Musique en jeu*, n. 17, p. 37–62, 1975.

MORIN, Edgar. A inteligência cega. In: MORIN, Edgar. *Introdução ao pensamento complexo*. 4. ed. Porto Alegre: Sulina, 2005a. p. 9–16.

MORIN, Edgar. A religação dos saberes: o desafio do século XXI. São Paulo: Bertrand Brasil, 2010.

MORIN, Edgar. *O método 3: O conhecimento do conhecimento*. 4. ed. Porto Alegre: Sulina, 2012.

MORIN, Edgar. O paradigma complexo. In: MORIN, Edgar. *Introdução ao pensamento complexo*. 4. ed. Porto Alegre: Sulina, 2005b. p. 57–78.

MURPHY, Timothy S.; SMITH, Daniel W. What I hear is thinking too: Deleuze and Guattari go pop. *ECHO: a music centered journal*, v. 2, 2001.

NAPOLITANO, Marcos. *História e Música: história cultural da música popular*. Belo Horizonte: Autêntica, 2002.

NATTIEZ, Jean-Jacques. O modelo tripartite de semiologia musical. *Debates - Cadernos do Programa de Pós-graduação em Música do Centro de Letras e Artes da UNIRIO*, n. 6, p. 7–39, 2003.

NESTROVSKI, Arthur. *Lendo música: 10 ensaios sobre 10 canções*. São Paulo: Publifolha, 2007.

OLSON, Hope A. Exclusivity, teleology and hierarchy: Our Aristotelean legacy. *Knowledge Organization*, v. 26, n. 2, p. 65–73, 1999.

ORTEGA, Cristina Dotta; LARA, Marilda Lopes Ginez De. A noção de documento: de Otlet aos dias de hoje. *Datagramazero - Revista de Ciência da Informação*, v. 11, n. 2, 2010.

ORTIZ, Renato. *A moderna tradição brasileira*. São Paulo: Brasiliense, 1988.

PATTUELLI, Cristina; RUBINOW, Sara. The knowledge organization of DBpedia: a case study. *Journal of Documentation*, v. 69, n. 6, p. 762–772, 2013.

PATTUELLI, M. Cristina; WELLER, Chris; SZABLYA, Genevieve. *Linked Jazz: an exploratory pilot*. 2011. p. 158–164.

PINHEIRO, Lena Vania Ribeiro; PEREIRA, Maria de Nazaré Freitas. O sonho de Otlet: aventura em tecnologia da informação e comunicação. *Interdiscursos da Ciência da Informação: arte, museu e imagem*, 2000.

PLEWA, Magdalena; KOSTEK, Bozena. Multidimensional Scaling Analysis Applied to Music Mood Recognition. *Audio Engineering Society*, 2013.

POLI, Roberto. A note on the difference between complicated and complex social systems. *Cadmus*, v. 2, n. 1, 2013.

POMBO, Olga. Da classificação dos seres à classificação dos saberes. *Revista da Biblioteca Nacional de Lisboa*, n. 2, p. 19, 1998.

RAIMOND, Yves *et al.* The Music Ontology. *ISMIR Proceedings*, 2007. p. 417–422.

RANGANATHAN, Shiyali Ramamrita. *Prolegomena to library classification*. 2. ed. Asia Publishing House (New York), 1967.

SADIE, Stanley. *The New Grove Dictionary of Music and Musicians*. London: Macmillan, 1980.

SAMPAIO, Luis Paulo. Reflexos no Espelho: Ensaio de Análise Semiológica Tripartite das variações op.27 de Anton Webern. *DEBATES - Cadernos do Programa de Pós-Graduação em Música*, n. 6, 2014.

SANTINI, Rose-Marie. *Os usuários e a desorganização da cultura: os sistemas de recomendação e as consequências da classificação para os usos sociais da música na internet*. Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, 2010.

SEMANTIC WEB. Disponível em: <<http://www.w3.org/standards/semanticweb/>>. Acesso em: mai 2016.

SEVERIANO, Jairo. *Getúlio Vargas e a música popular*. Rio de Janeiro: FGV-Instituto de Documentação, Editora da Fundação Getúlio Vargas, 1983.

SEVERIANO, Jairo; MELLO, Zuza Homem De. *A canção no tempo: 85 anos de músicas brasileiras*. 5. ed. São Paulo: Editora 34, 1997. v. 1.

SHERA, Jesse H. Putting knowledge to work. *Libraries and the Organization of Knowledge*, p. 51–62, 1956.

SHERA, Jesse H. What lies ahead in classification. *Proceedings of the Allerton Park Institute*, p. 116–128, 1960.

SOUZA, Rosali Fernandez de. A classificação como interface da internet. *Datagramazero - Revista de Ciência da Informação*, v. 2, n. 2, abr. 2000. Disponível em: <http://www.datagramazero.org.br/abr00/F_I_art.htm>. Acesso em: mai 2016.

SPOLADORE, Marina Carvalho. A aplicação de um modelo semiológico de análise em Savanas, de Almeida Prado. *Cadernos do Colóquio*, v. 9, n. 1, 2009.

STANECK, José; LANZELOTTE, Rosana Saldanha da Gama. Padrões de interoperabilidade para a difusão de recursos musicais. In: CONGRESSO DA ASSOCIAÇÃO NACIONAL DE PESQUISA E PÓS-GRADUAÇÃO EM MÚSICA - ANPPOM, 17, São Paulo, 2007.

STRACCIA, Umberto. Description logics with fuzzy concrete domains. In: CONFERENCE ON UNCERTAINTY IN ARTIFICIAL INTELLIGENCE, 21, Vancouver, 2012.

STRACCIA, Umberto. *Foundations of fuzzy logic and semantic web languages*. Flórida: CRC Press, 2013.

TANENBAUM, Andrew S. *Structured computer organization*. Londres: Pearson, 2006.

TATIT, Luiz. *Análise semiótica através das letras*. São Paulo: Ateliê Editorial, 2001.

TATIT, Luiz. *Musicando a semiótica: ensaios*. São Paulo: Annablume, 1998.

TATIT, Luiz. *O cancionista: composição de canções no Brasil*. São Paulo: Edusp, 1995.

TATIT, Luiz. *Semiótica da canção: melodia e letra*. São Paulo: Escuta, 1994.

THE TRADITIONAL SQUARE OF OPPOSITION. In: PARSONS, Terence. *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed.), 2015. Disponível em: <<http://plato.stanford.edu/archives/sum2015/entries/square>>. Acesso em mai 2016.

THOMPSON, Abbey E. *Playing tag: an analysis of vocabulary patterns and relationships within a popular music folksonomy*. Carolina do Norte: Chapel Hill, 2008.

TINHORÃO, José Ramos. *História social da música popular brasileira*. São Paulo: Editora 34, 1998.

VELMURUGAN, C. APUPA Pattern. *Library and Information Science Links*. 2010. Disponível em: <<http://lislinks.com/forum/topics/apupa-pattern>>. Acesso em: out 2015.

VIANNA, Hermano. *O mistério do samba*. Rio de Janeiro: Zahar, 1995.

VICKERY, Brian C. Faceted classification for the web. *Axiomathes*, v. 18, n. 2, p. 145–160, 2008.

VICKERY, Brian C. Information representation and retrieval in the digital age. *Program*, v. 38, n. 3, p. 213–213, 2004.

VICKERY, Brian C. Knowledge representation: a brief review. *Journal of Documentation*, v. 4, n. 3, p. 145–159, 1986.

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

WEBER, Leo; KLEIN, Pedro Antonio Trierweiler. *Aplicação da lógica fuzzy em software e hardware*. Canoas: ULBRA, 2003.

WEINGART, Scott B. *From trees to webs: uprooting knowledge through visualization*. 2013. p. 43–58.

WILDEN, Anthony. *System and Structure: essays in communication and exchange*, 2 ed. Abingdon: Routledge, 2013. v. 11.

WILLIAMS, Raymond. *Culture and society, 1780-1950*. Nova Iorque: Columbia University Press, 1983.

WISNIK, José Miguel. Getúlio da paixão cearense (Villa-Lobos e o Estado Novo). SQUEFF, Enio; WISNIK, José Miguel. *O nacional e o popular na cultura brasileira*, p. 129–191, 1982.

WISNIK, José Miguel. *Machado maxixe*. São Paulo: Publifolha, 2008.

WISNIK, José Miguel. *O som e o sentido: uma outra história das músicas*. São Paulo: Editora Companhia das Letras, 1989.

WISNIK, José Miguel. *Sem receita: ensaios e canções*. São Paulo: Publifolha, 2004.

YANG, Yi-Hsuan; LIU, Chia-Chu; CHEN, Homer H. Music emotion classification: a fuzzy approach. *ACM*, p. 81–84. 2006.

ZADEH, Lotfi A. Fuzzy languages and their relation to human and machine intelligence. In: INTERNATIONAL CONFERENCE ON MAN AND COMPUTER, 1971, Basel: S. Karger, p. 130–165, 1971.

ZADEH, Lotfi A. Fuzzy Sets. *Information and Control*, v. 8, n. 3, p. 338–353, 1965.

ZENG, Marcia L. Knowledge organization systems (KOS). *Knowledge Organization*, v. 35, n. 2–3, p. 160–182, 2008.

ZILBERBERG, Claude. As condições semióticas da mestiçagem. In: CAÑIZAL, E. P.; CAETANO, K. E. (Org.). *O olhar à deriva: mídia, significação e cultura*. Tradução de Ivã Carlos Lopes; Luiz Tatit. São Paulo: Annablume, 2004. p. 69–101.

ZILBERBERG, Claude. *Eléments de grammaire tensive*. Limoges: Presse Universitaire de Limoges, 2006a.

ZILBERBERG, Claude. *Razão e poética do sentido*. São Paulo: Editora da Universidade de São Paulo, 2006b.

ZILBERBERG, Claude. Relation et rationalité. Actualité de Brøndal. *Langages*, A l'occasion d'un centenaire: Actualité de Brøndal. v. 22, n. 86, p. 59–77, 1987.

REFERÊNCIAS DISCOGRÁFICAS

BARRO, João de [Braguinha]; ALMEIDA, Antonio. *A mulata é a tal*. Intérprete: Ruy Rey. In: Disco 78 RPM, número 15.841. Continental, 1948.

BARROSO, Ari. *Aquarela do Brasil*. Intérprete: Francisco Alves. In: Disco 78 RPM, número 11.768-a. Odeon, 1939.

BUARQUE, Chico. *Paratodos*. Intérprete: Chico Buarque. In: Paratodos. Long Play. Faixa 1. BMG Ariola, 1993.

JACKSON, Michael. *Black or white*. Intérprete: Michael Jackson. In: Dangerous. Compact Disc. Faixa 8. Sony Music, 1991.

MONTEIRO, Ciro. *Quatro loucos num samba*. Intérprete: Ciro Monteiro. In: Sr. Samba. Long Play. Faixa 1. Columbia, 1962.

TAPAJÓS, Maurício; BLANC, Aldir. *Querelas do Brasil*. Intérprete: Aldir Blanc e Maurício Tapajós. Long Play. Faixa 1. SACI (Sociedade de Artistas e Compositores Independente), 1984.

VELOSO, Caetano. *Americanos*. Intérprete: Caetano Veloso. In: Circuladô Vivo. Compact Disc. Faixa 2. Decca Records, 1992.

VELOSO, Caetano. *Sugar cane fields forever*. Intérprete: Caetano Veloso. In: Araçá Azul. Long Play. Faixa 7. Philips, 1973.