UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE ARTES E COMUNICAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

FELIPE MOZART DE SANTANA NASCIMENTO

SISTEMA DE ORGANIZAÇÃO DO CONHECIMENTO: um mapeamento de conceitos e processos para a Ciência da Informação no Brasil

RECIFE
2022

FELIPE MOZART DE SANTANA NASCIMENTO

SISTEMA DE ORGANIZAÇÃO DO CONHECIMENTO: um mapeamento de conceitos e processos para a Ciência da Informação no Brasil

Tese apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de Doutor em Ciência da Informação.

Área de concentração: Informação, Memória e Tecnologia.

Orientador: Prof. Dr. Fabio Assis Pinho

RECIFE

2022

Catalogação na fonte
Bibliotecária Jéssica Pereira de Oliveira – CRB-4/2223

N244s Nascimento, Felipe Mozart de Santana
Sistema de Organização do Conhecimento: um mapeamento de conceitos e processos para a Ciência da Informação no Brasil / Felipe Mozart de Santana Nascimento. – Recife, 2022.
273f.: il.

Sob orientação de Fabio Assis Pinho.
Tese (Doutorado) – Universidade Federal de Pernambuco. Centro de Artes e Comunicação. Programa de Pós-Graduação em Ciência da Informação, 2022.

Inclui referências.

1. Sistemas de Organização do Conhecimento. 2. Taxonomia. 3. Tesauro. 4. Ontologia. I. Pinho, Fabio Assis (Orientação). II. Título.

020 CDD (22. ed.) UFPE (CAC 2022-114)

FELIPE MOZART DE SANTANA NASCIMENTO

SISTEMA DE ORGANIZAÇÃO DO CONHECIMENTO: um mapeamento de conceitos e processos para a Ciência da Informação no Brasil

Tese apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de Doutor em Ciência da Informação.

Área de concentração: Informação, Memória e Tecnologia.

Data de aprovação: 01/02/2022

BANCA EXAMINADORA

_____
Prof. Dr. Fabio Assis Pinho (Orientador)
Universidade Federal de Pernambuco

_____
Prof. Dr. Fábio Mascarenhas e Silva (Examinador interno)
Universidade Federal de Pernambuco

_____
Prof.ª Dr.ª Majóry Karoline Fernandes de Oliveira Miranda (Examinador interno)
Universidade Federal de Pernambuco

_____
Prof.ª Dr.ª Luciane Paula Vital (Examinador externo)
Universidade Federal de Santa Catarina

_____
Prof.ª Dr.ª Zaira Regina Zafalon (Examinador externo)
Universidade Federal de São Carlos

Com o objetivo de fortalecer o ensino e pesquisa em Organização e Representação da Informação e do Conhecimento, dedico à toda comunidade de ensaístas, pesquisadores e profissionais da Ciência da Informação.

AGRADECIMENTOS

Pelas energias que me trouxeram até aqui, agradeço, primeiramente, à **Deus**. Pela fé e motivação que foram fontes inspiradoras desde a idealização deste documento.

À **Universidade Federal de Pernambuco** e ao **Programa de Pós-graduação em Ciência da Informação** por todas as oportunidades pedagógicas, científicas e técnicas que foram propostas durante meu caminhar acadêmico.

À **Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco** pelo financiamento da pesquisa durante os meses que acompanhou o desenvolvimento desta tese.

Ao meu orientador **Fábio Pinho** pela confiança, oportunidade, encorajamento e dedicação que conduziu todo o processo de orientação.

Aos **professores** que compuseram a banca de qualificação desta tese, em que muito contribuíram para o aperfeiçoamento da pesquisa.

À professora do Departamento de Ciência da Informação desta Universidade, **Márcia Braz**, que em muito contribuiu para o desenvolvimento da minha jornada profissional, abrindo as portas para dividir a sala de aula dos muitos estágios docência realizados.

Aos **alunos** das turmas de graduação e pós-graduação em que estive como docente. Vocês foram (e são), essencialmente, parte ativa e motivadora deste percurso.

Aos **eventos** profissionais e acadêmicos que favorecessem o compartilhamento de experiências, em especial ao Encontro Nacional de Pesquisa e Pós-graduação em Ciência da Informação (**ENANCIB**) e o Encontro Regional Norte-Nordeste de Educação em Ciência da Informação (**ERECIN**).

À minha **família** que, por muitos momentos, influenciaram-me a seguir a carreira que sonho desde criança e que alimentaram o sonho da pesquisa e desenvolvimento desta tese.

Aos **amigos** e **colegas** que estiveram ao meu lado nos momentos de angústias, de emoções, de felicidades etc.

A **todos** que contribuíram direta ou indiretamente para que esse percurso fosse trilhado com afinco.

A história dos registros do conhecimento e dos respectivos processos de tratamento, organização e recuperação, como tantas outras, ampliou-se e diversificou-se, mas não se apresenta assim de forma tão sistemática, linear e facilmente compreensível, como parece em uma primeira aproximação. Acrescentam-se aí novos e diversos suportes, processos, tecnologias, campos de conhecimento e profissionais, e suas diferentes combinações (SOUZA, 2015, p. 32).

RESUMO

O percurso histórico e científico da Ciência da Informação está orientado pelas práticas de tratamento e organização da informação e do conhecimento, bem como a representação, gestão, recuperação e acesso. Aliado a este percurso encontram-se os avanços das tecnologias digitais. A integração dessas práticas e o aperfeiçoamento dessas tecnologias motivaram o aprimoramento de conceitos; processos; instrumentos; métodos; e tantas outras operações inerentes aos ensaístas, pesquisadores e profissionais dessa ciência. Neste sentido, este estudo objetivou mapear os conceitos e processos dos Sistemas de Organização do Conhecimento (tesauro, taxonomia e ontologia) aperfeiçoados pelo aprimoramento das tecnologias digitais, para organização e representação da informação e do conhecimento na Ciência da Informação no Brasil. Promovendo, ainda, uma investigação qualitativa dos métodos de planejamento, construção e uso desses sistemas. Portanto, a tese estruturou-se em quatro seções: Introdução, em que se buscou a contextualização da pesquisa, delimitando o problema e a justificativa da investigação; a apresentação dos objetivos aos quais a pesquisa se propôs; o percurso metodológico adotado que, neste caso, apoiou-se na Metodologia da Problematização e no Arco de Maguerez. A segunda e terceira seções refletiram sobre a área de Organização e Representação da Informação e do Conhecimento, respectivamente, explorando, assim, um panorama bibliográfico sobre os sistemas alvo de estudo desta tese e outros, como: Classificação, Folksonomia e Mapas Conceituais. Em seguida, na seção de análise e discussão dos resultados, foram apresentadas as reflexões em que se propuseram os objetivos desta tese, mapeando, pois, um itinerário evolutivo do desenvolvimento de sistemas de organização do conhecimento, bem como verticalizando os estudos de conceitos e processos dos sistemas alvo de investigação desta pesquisa. Por meio dos elementos anteriormente mencionados, considerou-se que tesauros, taxonomias e ontologias, para a Ciência da Informação, em seu escopo conceitual-processual, foram potencialmente aperfeiçoados em razão das propriedades e do aprimoramento das tecnologias digitais adotadas no planejamento, execução, manutenção e avaliação da aplicação e do uso desses sistemas, nos múltiplos ambientes em que se apresentaram.

Palavras-chave: sistemas de organização do conhecimento; taxonomia; tesauro; ontologia.

ABSTRACT

The historical and scientific path of Information Science is guided by the practices of treatment and organization of information and knowledge, as well as the representation, management, retrieval and access. Allied to this path are the advances in digital technologies. The integration of these practices and the improvement of these technologies have motivated the improvement of concepts; processes; instruments; methods; and so many other operations related to essayists, researchers and professionals of this science. The aim of this thesis was to map the concepts and processes of the Knowledge Organization Systems (thesaurus, taxonomy and ontology) improved by the enhancement of digital technologies, for the organization and representation of information and knowledge in the Information Science in Brazil. It also promotes a qualitative investigation of the methods of planning, construction and use of these systems. Therefore, the thesis is structured in four sections: Introduction, in which the contextualization of the research was sought, delimiting the problem and the justification for the investigation; the presentation of the objectives to which the research was proposed; the methodological path adopted which, in this case, was based on the Problematization Methodology and on the Arch of Maguerez. The second and third sections reflected on the area of Organization and Representation of Information and Knowledge, respectively, exploring, thus, a bibliographical overview on the target systems studied in this thesis and others, such as: Classification, Folksonomy and Conceptual Maps. Then, in the section of analysis and discussion of results, the reflections in which the objectives of this thesis were proposed were presented, mapping, therefore, an evolutionary itinerary of the development of systems of knowledge organization, as well as verticalizing the studies of concepts and processes of the target systems of investigation of this research. Through the elements previously mentioned, it was considered that thesaurus, taxonomies and ontologies, for Information Science, in their conceptual-processual scope, were potentially improved due to the properties and the improvement of digital technologies adopted in the planning, execution, maintenance and evaluation of the application and use of these systems, in the multiple environments in which they were presented.

Keywords: knowledge organization systems; taxonomy; thesaurus; ontology.

LISTA DE FIGURAS

Figura 1 - Três perspectivas dedicadas à consolidação da Ciência da Informação ..22
Figura 2 - Arco de Maguerez.....27
Figura 3 - Fases da Análise de Conteúdo de Bardin.....31
Figura 4 - Características dos conceitos de informação e de conhecimento.....38
Figura 5 - Representação da Informação no processo de busca e de recuperação da informação.....42
Figura 6 - Perspectivas da Organização do Conhecimento e sua relação com a Ciência da Informação.....44
Figura 8 - Birger Hjorland 45
Figura 7 - Ingetraut Dahlberg 45
Figura 9 - Propriedades conceituais de Organização e Representação da Informação e do Conhecimento.....47
Figura 10 - Concepções conceituais de Organização e representação da informação e do conhecimento. 48
Figura 11 - Teoria do Conceito de Dahlberg: linguagem, objeto e relações.....51
Figura 12 - Teoria do Conceito de Dahlberg: objetos e conceitos.....52
Figura 13 - Teoria do Conceito de Dahlberg: relação lógica entre conceitos 53
Figura 14 - Teoria do Conceito de Dahlberg: relação hierárquica, partitiva, oposição e funcional entre conceitos 54
Figura 15 - Teoria do Conceito de Dahlberg: ordem, espécie, sentido e tipologia de conceitos 55
Figura 16 - Teoria do Conceito de Dahlberg: ordem das características dos conceitos 57
Figura 17 - Teoria do Conceito de Dahlberg: tipologias das características dos conceitos 58
Figura 18 - Teoria do Conceito de Dahlberg: sentido de intensão e extensão do conceito. 59
Figura 19 - Teoria do Conceito de Dahlberg: espécies de conceito 59
Figura 20 - As duas fases de SOC.....61
Figura 21 - Perspectivas conceituais de SOC. 62
Figura 22 - Exemplos de SOC quanto a estruturas e funções 64

Figura 23 - Contribuições dos SOC.....66
Figura 24 - Árvore de Porfírio.....70
Figura 25 - Aristóteles (384 – 322 a. C.).....71
Figura 26 - Platão (428 – 347 a. C.).....71
Figura 27 - As sete artes liberais.....73
Figura 28 - Processos para construção de Mapas Conceituais 83
Figura 29 - Tipos de folksonomia 85
Figura 30 - Processo cognitivo por trás da etiquetagem 86
Figura 31 - Uso de linguagem formalizada no processo de etiquetagem livre 87
Figura 32 - Uso de linguagem natural no processo de etiquetagem livre.....87
Figura 33 - Thesaurus of English Words. 90
Figura 35 - Plan of Classification of Thesaurus of English Words. 91
Figura 34 - Introduction of Thesaurus of English Words.....91
Figura 36 - Objetivos do sistema tesauro. 92
Figura 37 - Componentes do sistema de tesauro.....93
Figura 38 - Sistemas e subsistemas de tesauro.....94
Figura 39 - Exemplo de relação de equivalência.....95
Figura 40 - Exemplo de relação hierárquica.....96
Figura 41 - Exemplo de relação associativa.....97
Figura 42 - Exemplo (2) de relação associativa 98
Figura 43 - Exemplo de relacionamento ontológico 98
Figura 44 - Carlos Lineu.....100
Figura 45 - Systema Naturae 100
Figura 47 - Taxionomia de objetivos educacionais.....101
Figura 46 - Benjamim Bloom.....101
Figura 48 - Estrutura do sistema taxonomia.....104
Figura 49 - Aspectos relevantes sobre termos empregados na taxonomia. 105
Figura 50 - Integração dos conceitos de ontologia. 106
Figura 51 - Componentes de uma ontologia. 107
Figura 52 - Classificações de Ontologia.....109
Figura 53 - Itinerário evolutivo do desenvolvimento de SOC.....111
Figura 54 - Tecnologias para representação da informação em modelos interoperáveis.....114

Figura 55 - Processos da análise de dados e o papel do profissional da informação....117
Figura 56 - Concepções sobre tesauros....129
Figura 57 - Concepções estendidas sobre tesauros....130
Figura 58 - Características de SKOS....135
Figura 59 - Exemplos de tesauro....140
Figura 60 - Mapeamento dos processos para construção do tesauro por Valencia, Vogel e Silva....143
Figura 61 - Apresentação planigráfica....144
Figura 62 - Apresentação sistemática....145
Figura 63 - Apresentação alfabética....145
Figura 64 - Mapeamento dos processos para construção do tesauro por Maia, Sobrinho e Cundurú....147
Figura 65 - Apresentação de parte das relações do microtesauro sobre gestão ambiental....148
Figura 66 - Mapeamento dos processos para construção do tesauro por Vieira, Santos e Lapa.....149
Figura 67 - Mapeamento dos processos para construção do tesauro por Kroeff e Leoneti....150
Figura 68 - Representação em formato Tree do tesauro brasileiro de ciências do esporte....151
Figura 69 - Representação em formato Alpha do tesauro brasileiro de ciências do esporte....152
Figura 70 - Mapeamento dos processos para construção do tesauro por Kumar Bl e Nikam....153
Figura 71 - Dicionário de sinônimos de ioga com hiperlink de glossário de ioga em sânscrito-inglês Fonte: Kumar Bl e Nikam (2014, p. 292). 154
Figura 72 - Mapeamento dos processos para construção do tesauro por Souza *et al*. 155
Figura 73 - Tesauro Phytoplankton 157
Figura 74 - Mapeamento dos processos para construção do tesauro por Rosati *et al.* 158

Figura 75 - Mapeamento dos processos para construção do tesauro por Tartarotti, Dal'Evedove e Fuijta.....160
Figura 76 - Mapeamento dos processos para construção do tesauro por Elena Víquez.....162
Figura 77 - Mapeamento dos processos para construção do tesauro por Nancy Orrego.....164
Figura 78 - Mapeamento dos processos para conversão de tesauro estático para um software gestor de tesauro por Santos, Cervantes e Fujita.....167
Figura 79 - Codificação em PHP do algoritmo que gera o arquivo etiquetado de importação; e a estrutura sistemática do TBCI.....170
Figura 80 - Ferramentas para construção de tesauros.....173
Figura 81 - Concepções sobre taxonomia.....185
Figura 82 - Concepções estendidas sobre taxonomia.....186
Figura 83 - Interseção entre Taxonomia e Organização, Gestão e Tecnologia.....188
Figura 84 - Ecologia da Informação e o Arquiteto de Informação.....192
Figura 85 - Exemplos de taxonomia.....193
Figura 86 - Mapeamento dos processos para construção da taxonomia por Vital e Café.....195
Figura 87 - Mapeamento dos processos para construção da taxonomia por Aganette e Teixeira.....197
Figura 88 - Figura 88 –Mapeamento dos processos para construção da taxonomia por Bernardino, Simões, Gracioso e Costa.....198
Figura 89 - Mapeamento dos processos para construção da taxonomia por Wu e Yang.....199
Figura 90 - Esquema lógico do uso de taxonomia da relação do derramamento de óleo para efetuar inferências.....200
Figura 91 - Mapeamento dos processos para construção da taxonomia por Fitterer, Mettler, Rohner e Winter.....201
Figura 92 - Mapeamento dos processos para construção da taxonomia por Silva, Freitas e Silva.....203
Figura 93 - Mapeamento dos processos para construção da taxonomia por Braga.....205
Figura 94 - Concepções sobre ontologia.....216

Figura 95 - Concepções estendidas sobre ontologia 217
Figura 96 - Exemplos de ontologia 221
Figura 97 - Mapeamento dos processos para construção da ontologia por Mendonça e Almeida 224
Figura 98 - Mapeamento dos processos para construção da ontologia por Dziekaniak 229
Figura 99 - Seções do Ontokem 230
Figura 100 - Domínio e range em ontologia 231
Figura 101 - Mapeamento dos processos para construção da ontologia por Sousa et al 232
Figura 102 - Mapeamento dos processos para construção da ontologia por Rautenberg, Todesco e Steil 237
Figura 103 - Mapeamento dos processos para construção da ontologia por Ramalho, Zafalon e Ouchi 238
Figura 104 - Ferramentas para aprendizagem de ontologias 242

LISTA DE QUADROS

Quadro 1 - Caracterização do itinerário metodológico 28
Quadro 2 - Fontes bibliográficas 29
Quadro 3 - Síntese das dimensões de Rafael Capurro. 35
Quadro 4 - Sistemas Escolásticos de Organização do Conhecimento e Sistemas Científicos de Organização do Conhecimento 75
Quadro 5 - A notação e os Sistemas bibliográficos de organização do conhecimento 76
Quadro 6 - Sistemas de organização do conhecimento bibliográficos 77
Quadro 7 - Concepções conceituais sobre tesauro. 118
Quadro 8 - Propriedades SKOS. 136
Quadro 9 - Exemplos de relações semânticas dos tesauros e Thesa. 138
Quadro 10 – Concepções conceituais sobre taxonomia. 176
Quadro 11 – Concepções conceituais sobre ontologia. 208

LISTA DE SIGLAS

ANCIB	Associação Nacional de Pesquisas e Pós-graduação em Ciência da Informação
BRAPCI	Base de Dados em Ciência da Informação
CC	Ciência da Computação
CI	Ciência da Informação
ENANCIB	Encontro Nacional de Pesquisas e Pós-graduação em Ciência da Informação
ERECIN	Encontro Regional Norte Nordeste de Educação em Ciência da Informação
ISKO	*International Society for Knowledge Organization*
KOS	*Knowledge Organization System*
NASKO	*North American Symposium on Knowledge Organziation*
OC	Organização do Conhecimento
OI	Organização da Informação
OIC	Organização da Informação e do Conhecimento
ORIC	Organização e Representação da Informação e do Conhecimento
SOC	Sistemas de Organização do Conhecimento
SRI	Sistema de Recuperação da Informação
TI	Tecnologia da Informação

SUMÁRIO

1 INTRODUÇÃO 19
1.1 Contextualização da pesquisa: problema e justificativa 20
1.2 Objetivos 26
1.3 Percurso Metodológico 26
1.4 Estrutura da pesquisa 32
2 ORGANIZAÇÃO E REPRESENTAÇÃO DA INFORMAÇÃO E DO CONHECIMENTO 34
2.1 Informação e Conhecimento: breves perspectivas conceituais 34
2.2 Organização e Representação: da informação ao conhecimento 39
2.3 Organização e Representação da Informação 40
2.4 Organização e Representação do Conhecimento 43
2.5 A Teoria do Conceito de Dahlberg 49
3 SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO 61
3.1 Conceitos e características de SOC 61
3.2 Classificações 67
3.3 Mapas Conceituais 80
3.4 Folksonomia 83
3.5 Tesauro 89
3.6 Taxonomia 99
3.7 Ontologia 105
4 ANÁLISE E DISCUSSÃO DOS RESULTADOS 111
4.1 Itinerário evolutivo do desenvolvimento de SOC: interoperabilidade e big data 111
4.2 Tesauros: conceitos e processos 118
4.3 Taxonomia: conceitos e processos 175
4.4 Ontologia: conceitos e processos 207
5 CONSIDERAÇÕES FINAIS 246
REFERÊNCIAS 252

1 INTRODUÇÃO

Os aspectos epistemológicos, que orientam os estudos voltados à problemática informacional, se consolidaram há pouco. A idealização de uma ciência que abarcasse os estudos voltados à informação foi constituída por outras disciplinas que possuíam propriedades e finalidades distintas. Essas disciplinas contribuíram para o processo de consolidação científica da Ciência da Informação.

A consolidação da Ciência da Informação reúne funcionalidades e atributos de três disciplinas que antecederam sua legitimação, isto é: Bibliografia, Documentação e Biblioteconomia. Das quais, a primeira surgiu a partir da criação da imprensa e do volume de livros produzidos na Europa, tendo como contributo social e científico reunir a produção intelectual humana, que se encontrava materializada em livros e manuscritos distribuídos em bibliotecas. A Documentação, por sua vez, buscava listar os registros dos ambientes em que eram depositadas as produções - essa listagem foi vista como um trabalho cooperativo entre instituições. A Biblioteconomia, em primeiro momento, possuía caraterísticas advindas da Bibliografia e da Documentação, mantendo relações funcionais e conceituais, mas diferenciando-se pelo tratamento e pela organização da informação, ou seja, os atributos técnicos (ARAÚJO, 2018).

Além das contribuições das linhas apresentadas acima, outros fatores foram determinantes para a sustentação da Ciência da Informação como ciência, a exemplo da teoria matemática da comunicação, proposta por Shannon e Weaver, e da teoria sistêmica de Bertalanffy e Wiener. Shannon e Weaver se fundamentaram no dinamismo da transmissão de mensagens, buscando representar o processo de comunicação, enquanto a teoria de Bertalanffy e Wiener representa a ideia de totalidade, mostrando que o total é o produto das partes que o compõem. (ARAÚJO, 2018; SOUZA, 2015).

Seja por meio de disciplinas, como a Bibliografia, a Documentação e a Biblioteconomia, ou de teorias impulsionadas por outros campos científicos, como a teoria matemática da comunicação, mencionada anteriormente, que foi formulada por estudos da Física e da Estatística, e a teoria sistêmica inspirada na Biologia, a CI buscava, em princípio, agregar os principais eixos científicos e técnicos, para ser vista como uma ciência que, então, possuía condições e estratégias suficientes para

se firmar como um campo científico, assumindo o compromisso de solucionar a problemática informacional existente.

A tecnologia da informação e da comunicação, em que muito contribuiu para que a CI fosse vista como uma nova ciência, é tida como elemento-chave por acompanhar a linha evolutiva da Ciência da Informação desde as primeiras discussões para sua legitimação científica. Em cada teoria que fundamenta a CI, pode-se apontar contribuições tecnológicas, sejam elas orientadas ao tratamento técnico, como organização e representação, ou às atividades de recreação, como produção, disseminação e uso da informação (ARAÚJO, 2018).

Nesse sentido, estabelece-se a ideia de que a Ciência da Informação, em sua essência, assume um caráter interdisciplinar, ora reunindo disciplinas científicas distintas, ora assumindo teorias originadas em outras áreas, para percorrer seu transcurso científico.

A CI está pautada no estabelecimento de subáreas, entre as décadas de 60 e 90, que fortaleceram sua ascensão científica. De acordo com Araújo (2018) tais subáreas são: Comunicação Científica; Organização do Conhecimento; Comportamento Informacional; Gestão da Informação e do Conhecimento; Economia Política da Informação; Estudos Métricos da Informação; e estão relacionadas, também, ao protagonismo tecnológico, que, acompanha o exercício informacional desempenhado por cada uma delas.

Considerando o que é dito por Araújo (2018) esta tese se concentra na “Organização do Conhecimento”, contextualizando os aportes científicos acompanhados do protagonismo tecnológico e apresentando um delineamento histórico-epistêmico dessa subárea, considerada, aqui, como campo científico.

A subseção a seguir busca apresentar elementos motivacionais, que despertaram interesse pelo desenvolvimento desta pesquisa, apontando os principais pontos a serem tratados nesta tese.

1.1 Contextualização da pesquisa: problema e justificativa

A compreensão dos processos históricos que levaram à constituição da CI é imprescindível para o desenvolvimento desta pesquisa, pois necessita-se de uma percepção linear acerca do constructo desta área do conhecimento científico.

Conforme apresentado na seção anterior, a CI integra subáreas que contribuíram para sua consolidação e legitimação como ciência; dentre as quais, esta pesquisa aprofundar-se-á em discussões voltadas ao campo da organização da informação e do conhecimento.

Na Ciência da Informação, “[...] a história dos respectivos processos de tratamento, organização e recuperação, como tantas outras, ampliou-se e diversificou-se” (SOUZA, 2015, p. 32). Isto acontece porque a sociedade está sempre em busca de novas formas para tratar e organizar a informação, por métodos tanto manuais quanto tecnológicos. Assim, a forma de tratamento e organização se modificou, sobretudo em virtude da introdução dos elementos tecnológicos. Para Souza (2015, p. 32), na linha histórica da Ciência da Informação, acrescentam-se “[...] novos e diversos suportes, processos, tecnologias, campos de conhecimento e profissionais, e suas diferentes combinações”.

Esse percurso científico transformador está evidenciado nas três bases epistemológicas da Ciência da Informação, estando elas apoiadas nas três grandes disciplinas que contribuíram para a concepção da CI, isto é, Bibliografia, Documentação e Biblioteconomia, e caracterizadas, também, pela tecnologia, conforme a apresentação da Figura 1.

Figura 1 - Três perspectivas dedicadas à consolidação da Ciência da Informação

Fonte: O autor (2022) baseado em Souza (2015).

A primeira perspectiva, **“**Bibliográfica Documental”, preocupa-se com o conjunto de sistemas, fundamentos e técnicas de organização da informação, que se constituem em embriões das tecnologias; a segunda é a “Biblioteconômica”, e debruça-se sobre as práticas de organização e tratamento da informação, apoiando-se em suportes que facilitem o acesso à informação pelos usuários (SOUZA, 2015); a terceira “Tecnológica”, foi provocada por um

> [...] deslocamento das discussões em torno das práticas bibliotecárias e documentárias propriamente ditas para um conjunto de atividades mais diretamente relacionadas ao planejamento e à organização de modernos sistemas de informação. (SOUZA, 2015, p. 54)

Portanto a preocupação desta perspectiva está centrada nas estratégias de recuperação da informação.

Um elemento evidente nas três perspectivas é a organização da informação, mesmo que orientada à recuperação e do acesso às informações. Conforme Araújo (2018), as primeiras manifestações desse elemento, na CI, surgiram devido ao

desafio de buscar a melhor forma de representar a informação, pensando na otimização da sua recuperação. Logo, uma das releituras desta pesquisa está centrada nos sistemas informacionais, que objetivam, além da recuperação e do acesso[1] às informações, a organização e a representação dos processos informacional e comunicacional aportando-se de princípios tecnológicos.

Nesse sentido, de compreender uma das releituras promovidas por esta tese, corrobora-se com Greenberg ao refletir que

> Hoje, há pouco mais de uma década no século 21, a tecnologia pode ser usada para quebrar as barreiras tradicionais da organização do conhecimento. Nesse campo, há muitos projetos explorando meios de compartilhar, manipular, estender e avançar nossa compreensão teórica e o uso da organização do conhecimento. (GREENBERG, 2011, p. 13, tradução nossa)

De acordo com Jorente, Padua e Santarem Segundo (2017) a complexidade dos sistemas, que cercam o processo informacional e comunicacional, visa o relacionamento entre tais elementos e suas partes. Assim, percebe-se uma convergência de informações nas estruturas sistêmicas complexas, que fazem emergir novos conceitos representacionais da informação. Como exemplo dessa atualização, no campo da organização e representação da informação e do conhecimento, são percebidas atualizações dos Sistemas de Organização do Conhecimento por meio do aprimoramento das tecnologias digitais, que, por sua vez, cooperaram para o aperfeiçoamento de conceitos e processos, e interferem, através da informação, nas interações dos indivíduos, criando cognições e um redesenho dos sistemas existentes.

Jiménez (2016) aponta uma outra evidência: as mudanças ocorridas nas formas de produção e difusão da informação e do conhecimento, às quais afetam técnicas e finalidades de organização e representação do conhecimento. Os SOC, no entanto, podem ser considerados como exemplo dessas mudanças.

Destaca-se, aqui, um retorno à Greenberg (2011, p. 14, tradução nossa) ao apontar que explorar pesquisas sobre SOC é cada vez mais relevante. Essa relevância é dada ao crescimento dos projetos de estruturação do conhecimento

[1] Para aqueles que precisam de informação o que interessa, em última análise, é ter *acesso* à informação desejada. Por isso, pode-se dizer que o acesso à informação é a questão básica de interesse da Ciência da Informação, em todas as suas especialidades (Biblioteconomia, Documentação, Arquivologia, Gestão da Informação etc.) e modalidades (pesquisa básica, pesquisa aplicada, tecnologia, economia, política etc.) (DIAS; NAVES, 2012, p. 2).

semântico, classificatório e ontológico. Logo, o papel desta tese é o de mapear as atualizações dos sistemas tesauro, taxonomia e ontologia, aperfeiçoados em conceitos e processos causados pelo aprimoramento das tecnologias digitais.

O desenvolvimento desta pesquisa justifica-se pelo atual cenário, social, cognitivo e tecnológico, que, em constantes mudanças, evidencia a necessidade de estudos teóricos e empíricos sobre os Sistemas de Organização do Conhecimento. Assim, uma pesquisa sobre as possíveis convergências e/ou divergências de conceitos e processos na construção e uso desses sistemas, para a Ciência da Informação, faz-se necessária. Essa necessidade evidencia-se, também, em Greenberg (2011, p. 14, tradução nossa) por considerar a organização do conhecimento como elemento vital e fundamental para a Ciência da Informação e Biblioteconomia. Acrescenta-se ainda aqueles outros campos como Gestão da Informação, Arquivologia e Documentação.

Este trabalho justifica-se, também, pela questão das representações informacionais modificadas pelos sistemas, que são veiculadas nas plataformas web e são usadas como um meio de compartilhamento e disseminação de informação/conhecimento, criando assim modelos de percepção de contato dos indivíduos e cognição dos agentes, sejam humanos ou máquinas (JORENTE; PADUA; SANTAREM SEGUNDO, 2017).

Acrescenta-se ainda que “[...] imersos nas tarefas diárias, raramente os estudamos – SOC, de uma perspectiva teórica que nos permite vislumbrar sua natureza, função e importância” (SUÁREZ SÁNCHEZ, 2017, p. 3). É, neste sentido, que esta pesquisa toma por base, a identificação dos aspectos contribuintes desses sistemas para o “exercício diário” de profissionais bibliotecários, arquivistas, documentalistas, gestores e cientistas da informação. Assim, refletir sobre SOC é de forte relevância, pois quanto mais sabe-se sobre eles, pode-se selecionar e implementar os mais adequados para resolução de problemáticas informacionais.

Destarte, os Sistemas de Organização do Conhecimento são estruturas sistemáticas, em que se percebem a construção de entidades abstratas da realidade, representando os conceitos de um determinado domínio, e a utilização desses sistemas para tratamento da informação e do conhecimento, tendo como consequência a Recuperação da Informação (RI) em ambientes distintos, cujas práticas estão diretamente ligadas aos métodos, técnicas e instrumentos de RI.

Neste sentido, como questão norteadora desta pesquisa tem-se *como os conceitos e processos dos Sistemas de Organização do Conhecimento (tesauro, taxonomia e ontologia) foram aperfeiçoados em relação ao aprimoramento das tecnologias digitais?* Utilizando, para tanto, o campo de organização e representação da informação e do conhecimento no Brasil.

Segundo a concepção individual do autor desta pesquisa, deve-se levar em consideração a trajetória acadêmica que vislumbra produções científicas relevantes para a CI, desenvolvidas sob a ótica dos Sistemas de Organização do Conhecimento, sendo maior ênfase as ontologias. Em perspectiva social, esta pesquisa trabalha de forma cooperativa com a Ciência da Informação, contribuindo para um dos objetivos principais desta ciência: o tratamento e a organização sistemática da informação e do conhecimento.

Em caráter de relevância científica, por meio da reunião das reflexões acima, o ineditismo e originalidade desta pesquisa centra-se no mapeamento de características, conceitos e processos que os sistemas de organização do conhecimento apresentam à comunidade da Ciência da Informação em seus aspectos teóricos, técnicos e práticos; além da significativa contribuição ao campo da organização e representação da informação e do conhecimento ao promover uma pesquisa orientada aos estudos sobre a construção desses sistemas. A reunião dessas propriedades torna essa pesquisa de forte relevância para a CI no Brasil, por promover um estudo que abrange boas práticas sobre SOC.

Por meio dos quesitos introdutórios, compreende-se que este trabalho se concentra na linha de pesquisa “Memória da Informação Científica e Tecnológica” do Programa de Pós-graduação em Ciência da Informação da Universidade Federal de Pernambuco[2], por ela preocupar-se com questões teóricas, conceituais e reflexivas, além de abordar pesquisas voltadas à Organização e à Representação da Informação e do Conhecimento.

Ademais, este trabalho dialoga com o grupo de pesquisa “Organização e Representação do Conhecimento”[3], que é certificado pelo Conselho Nacional de

[2] https://www.ufpe.br/ppgci/o-programa

[3] O grupo de pesquisa articula-se com as tendências na área de Organização do Conhecimento, em especial as discussões apontadas no Grupo de Trabalho 2 da Associação Nacional de Pesquisa e Pós-graduação em Ciência da Informação e com o capítulo brasileiro da International Society for Knowledge Organization (ISKO).

Desenvolvimento Científico e Tecnológico (CNPq)[4] e liderado pelos professores Fabio Assis Pinho e Marivalde Moacir Francelim, do qual o autor desta tese faz parte, integrando, portanto, a linha de pesquisa “Organização do Conhecimento e as Tecnologias de Informação”[5].

A contextualização científica deste estudo concentra-se para além desta seção, na apresentação dos objetivos que sustentam o desenvolvimento desta pesquisa. Esses objetivos estão apresentados a seguir.

1.2 Objetivos

Esta pesquisa tem como **objetivo geral** mapear os conceitos e processos dos Sistemas de Organização do Conhecimento (tesauro, taxonomia e ontologia) aperfeiçoados pelo aprimoramento das tecnologias digitais, para organização e representação da informação e do conhecimento na Ciência da Informação no Brasil.

E os **objetivos específicos** são quatro, a saber:

a) Reunir e caracterizar SOC;
b) Identificar possíveis convergências e/ou divergências entre os conceitos dos SOC (tesauro, taxonomia e ontologia) destacando contribuições tecnológicas digitais;
c) Identificar possíveis convergências e/ou divergências entre os processos dos SOC (tesauro, taxonomia e ontologia) destacando contribuições tecnológicas digitais;
d) Sistematizar processos para construção de tesauro, taxonomia e ontologia.

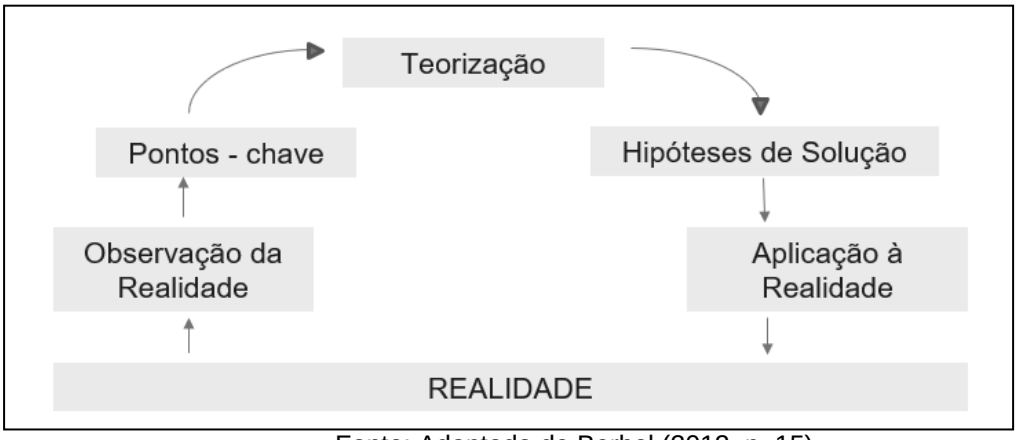
1.3 Percurso Metodológico

Dos aspectos relacionados à classificação, de acordo com Baffi (2002, p. 1), esta pesquisa pode ser considerada metodológica, pois está “[...] voltada para a inquirição de métodos e procedimentos adotados como científicos”, e teórica, por estar “[...] orientada no sentido de reconstruir teorias, quadros de referência, condições explicativas da realidade, polêmicas e discussões pertinentes”.

[4] http://dgp.cnpq.br/dgp/espelhogrupo/5524528636611242
[5] http://dgp.cnpq.br/dgp/espelhorh/0065600320

Vale ressaltar que a caracterização do percurso metodológico, adotada nesta pesquisa, está estritamente apoiada no objetivo geral elencado anteriormente. Protagoniza-se nesta tese a aplicação do Arco de Maguerez e Metodologia da Problematização. As etapas que formam o arco estão apresentadas na figura abaixo.

Figura 2 - Arco de Maguerez



Fonte: Adaptado de Berbel (2012, p. 15).

De acordo com Berbel (2012, p. 15) a problematização promove análises e críticas da realidade, uma vez que:

> [...] as ações para o desenvolvimento do processo partem da observação da realidade associado à temática eleita para o estudo. Da observação analítica e crítica dessa parcela da realidade é extraído um problema relevante para o estudo. Seguem-se as definições dos aspectos do problema a estudar, o estudo propriamente dito de tais aspectos, que preparam para as últimas etapas, das hipóteses de solução e da aplicação de uma ou mais dessas hipóteses na realidade da qual se extraiu o problema.

Neste sentido, delimitaram-se os seguintes passos trabalhados:

A “Observação da Realidade” concentrou-se no objetivo específico (a) de reunir e caracterizar SOC, com o intuito de refletir sobre a percepção do conceito de SOC e caracterizá-los quanto a estrutura e funções; este objetivo está evidenciado na seção 3 – Sistemas de Organização do Conhecimento.

Os “Pontos-Chave” implicaram no recorte dos sistemas que são sujeitos de estudos nesta pesquisa, são eles: tesauro, taxonomia e ontologia.

Os objetivos específicos voltados à “Teorização” concentraram-se em (b) e (c) que, juntos, buscam *identificar possíveis convergências e/ou divergências entre conceitos e processos dos SOC, destacando contribuições tecnológicas digitais.*

Para alcançá-los, foram estudadas as produções científicas dos campos que exploram esses sistemas, com a finalidade de encontrar os escopos conceituais que cercam o tema e suas distintas visões conceituais e processuais; cumpre-se, portanto, os objetivos b) e c) na seção 4 - “Análise e discussão dos resultados”.

As “Hipóteses de Solução” estão ligadas ao reconhecimento de como essas ideias se configuram hoje, permeando os elementos conceituais e processuais.

No que tange a “Aplicação da Realidade” esta dirige-se ao objetivo específico (d) *sistematizar processos para construção de tesauro, taxonomia e ontologia*. A fim de cumpri-lo, será explorado o estudo verticalizado sobre o funcionamento desses sistemas, desde os momentos de planejamento, construção e uso, ressaltando a necessidade de apresentar um novo olhar, sobre o cenário empírico-teórico dos SOC à comunidade da CI no Brasil; a sistematização proposta no objetivo específico d) contempla-se na seção 4 – “Análise e discussão dos resultados” e está apresentada com múltiplos percursos metodológicos modelados pela ferramenta Bizagi. Essas modelagens objetivam elucidar as ideias dos autores na execução e operação de taxonomias, tesauros e ontologias.

Compreende-se, portanto, que a integração dos objetivos específicos caracterizados anteriormente, por meio da aplicação da Metodologia da Problematização e o Arco de Maguerez, resultou no alcance dos resultados que a pesquisa propôs atingir.

Como apoio ao percurso metodológico, o Quadro 1, apresentado abaixo exemplifica os objetivos específicos e os processos adotados para alcançá-los.

Quadro 1 **-** Caracterização do itinerário metodológico

OBJETIVOS ESPECÍFICOS	APLICAÇÃO NO ARCO DE MAGUEREZ	COLETA DE DADOS	ANÁLISE DOS DADOS
I) Reunir e caracterizar SOC	Observação da realidade e pontos-chave	Pesquisa bibliográfica	Análise de conteúdo
II) Identificar possíveis convergências e/ou divergências entre os conceitos dos SOC destacando contribuições tecnológicas digitais.	Teorização e hipóteses de solução	Pesquisa bibliográfica	Análise de conteúdo

III) Identificar possíveis convergências e/ou divergências entre os processos dos SOC destacando contribuições tecnológicas digitais.	Teorização e hipóteses de solução	Pesquisa bibliográfica	Análise de conteúdo
IV) Sistematizar processos para construção de tesauro, taxonomia e ontologia.	Aplicação da realidade	Dados da pesquisa	Interpretação e inferência da pesquisa.

Fonte: Adaptado de Vital (2017, p. 118).

Considerando o Quadro 1 e as classificações em que se enquadram esta tese, tanto pesquisa teórica e metodológica quanto aos fins, concebe-se o processo de coleta de dados está apoiado em pesquisas bibliográficas em fontes reconhecidas nacional e internacionalmente pela CI, adotando como princípio de janela temporal o intervalo dos anos 2010 à 2019, com exceção das publicações dispostas nos anais dos Capítulos ISKO Brasil e NASKO Estado Unidos e Canadá, em que, respectivamente, o primeiro pauta-se no intervalo de 2012 à 2019 e o segundo no intervalo de 2007 à 2019. Cabe ressaltar que na tipologia “Anais” a busca foi realizada no arquivo ou página de cada evento, e nas “estratégias de busca” em fontes internacionais foi considerado o idioma dominante da língua local.

As fontes bibliográficas desta tese estão baseadas em publicações nacionais e internacionais que contemplam estudos relacionados ao campo de organização e representação da informação do conhecimento, tal como pode-se observar no Quadro 2.

Quadro 2 - Fontes bibliográficas

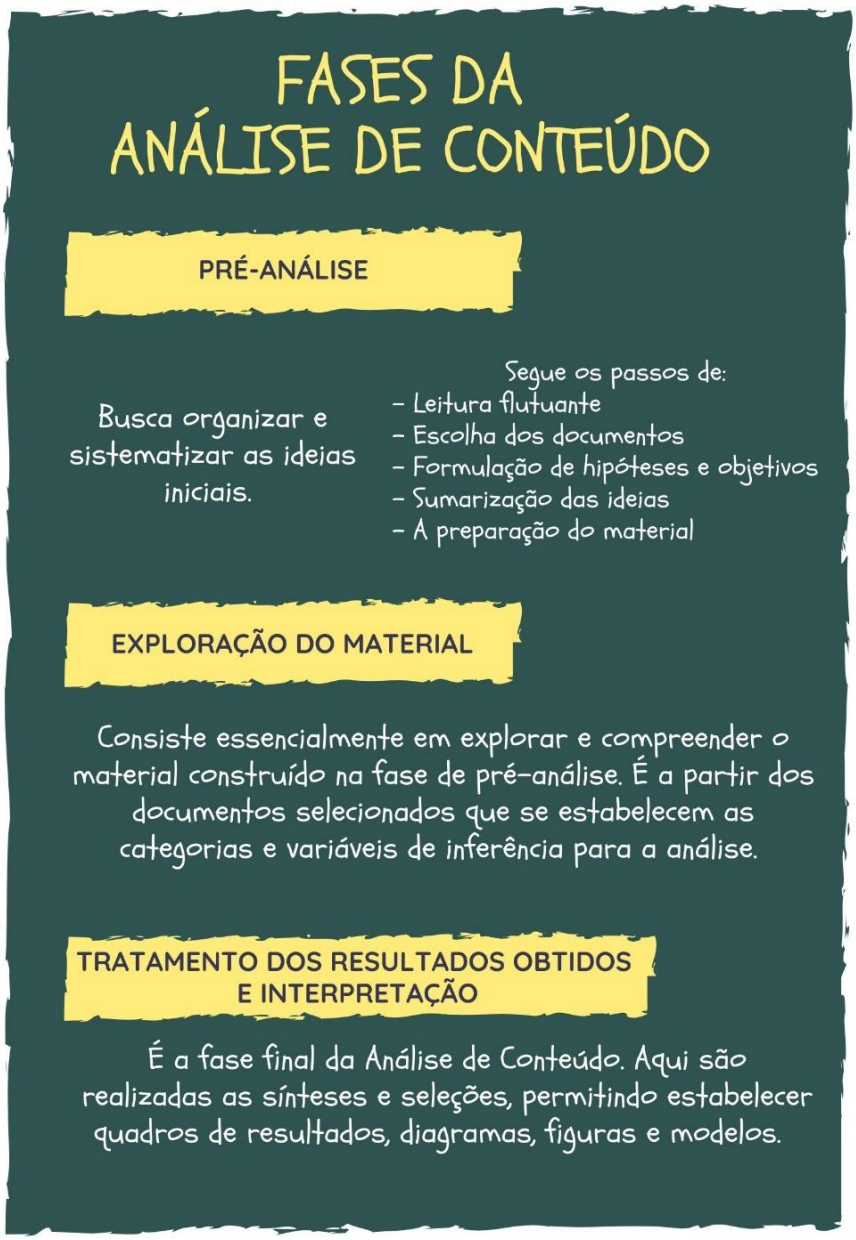
TIPOLOGIAS	FONTES	ESTRATÉGIA DE BUSCAS
BASE DE DADOS	**Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação (BRAPCI)**	**Título, resumo e palavras-chave (2010-2019).** “tesauro”; “taxonomia”; “ontologia”.
	Web of Science	**Título, resumo e palavras-chave (2010-2019).** “Thesauri AND information science AND knowledge organization system”; “Taxonomy AND information science AND knowledge organization system”; “Ontology AND information science AND

		knowledge organization system".
PERIÓDICO CIENTÍFICO	**Scire**	**Título, resumo e palavras-chave.** "thesauri"; "tesauro"; "thesaurus"; "taxonomia"; "ontologia".
ANAIS	**Anais do Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB)**	**Busca no título.** "Tesauro"; "taxonomia"; "ontologia".
	Capítulos ISKO: **ISKO-Brasil e NASKO-Canadá e Estados Unidos**	**Busca no título atendendo os eventos** Brasil (2012 a 2019); Canadá e Estados Unidos (2007 a 2019).

Fonte: O autor (2022).

Salienta-se que os “resultados” passaram pela análise de coleta de dados que compõe as fases da Análise de Conteúdo, de Bardin (2011). Conforme a autora, essas fases, caracterizadas na Figura 3, são: (a) pré-análise, (b) exploração do material; e (c) tratamento dos resultados, inferência e interpretação.

Figura 3 - Fases da Análise de Conteúdo de Bardin



Fonte: O autor (2022) com base em Bardin (2011) e Vital (2017).

A aplicação da análise de conteúdo na coleta bibliográfica considerou três fases: *Fase I* corresponde ao total de produções recuperadas; a *Fase II* corresponde à análise das produções por meio da leitura dos resumos, considerando aquelas que

abordam as temáticas desta tese; a *Fase III* apresenta a exploração da leitura integral da produção considerando a abordagem temática e as especificidades acerca de tesauro, taxonomia e ontologia, considerando os objetivos deste estudo.

As práticas de Organização do Conhecimento buscam a sistematização dos meios, para facilitar o processo de recuperação de um conhecimento já registrado e socializado, promovendo o acesso, o uso e a apropriação da informação pelos usuários. Disto isso, estabelece-se que os SOC devem ser tema de investigação constante, principalmente por desempenharem uma função de integração entre informação, conhecimento, e as fronteiras socioculturais e tecnológicas (BEGHTOL, 2005; CASARIM; FERREIRA; MILANI, 2017).

No sentido de aproximar e de relacionar os aspectos metodológicos ao desenvolvimento desta pesquisa, encontra-se, a seguir, a estrutura em que esta tese se apresenta.

1.4 Estrutura da pesquisa

Após a contextualização do estudo, a pesquisa estrutura-se da seguinte forma:

Na seção 2, intitulada “Organização e Representação da Informação e do Conhecimento”, são abordadas perspectivas teóricas da organização e representação da informação e do conhecimento, bem como caracterização da Teoria do Conceito, a qual sustenta o desenvolvimento de Sistemas de Organização do Conhecimento.

Na seção 3, intitulada “Sistemas de Organização do Conhecimento”, estão compreendidos aspectos basilares sobre os SOC, quanto a reunião e caraterização de SOC, além de serem apresentados: sistemas de organização do conhecimento bibliográfico, mapas conceituais e folksonomia. Esses SOC aparecem para que os outros tipos de sistemas, além dos que são alvo de estudo desta pesquisa, sejam expostos à comunidade.

Na seção 4, intitulada “Análise e discussão dos resultados” apresenta uma sistematização dos resultados da pesquisa. Está contextualizada pelo itinerário dos objetivos específicos, seguidos pelos sistemas tesauro, taxonomia e ontologia.

Em sequência são apresentadas as considerações finais, às quais compreendem as análises gerais desta tese, e as referências utilizadas para o embasamento da pesquisa.

2 ORGANIZAÇÃO E REPRESENTAÇÃO DA INFORMAÇÃO E DO CONHECIMENTO

Esta seção apresenta perspectivas conceituais sobre: informação e conhecimento; organização e representação da informação e do conhecimento; e caraterização da Teoria do Conceito de Dahlberg.

2.1 Informação e Conhecimento: breves perspectivas conceituais

As fronteiras, que delimitam os conceitos de informação e conhecimento, são temas recorrentes de pesquisas no âmbito da CI, seja por meio do ensino, da pesquisa ou da prática. As ideias discutidas nesta subseção não objetivam esgotar o mapeamento conceitual dos termos em questão, mas caracterizá-los a partir dos teóricos utilizados.

Na introdução desta tese foram citadas as disciplinas que deram origem à CI, bem como as teorias que contribuíram para sua consolidação. Foi neste sentido que autores se dedicaram à investigação e ao estabelecimento dos conceitos de informação, considerados objeto de estudo da CI. Araújo (2018) resgata histórica e geograficamente vários autores, que investigaram e definiram os conceitos relativos ao termo “informação”. Sucintamente, Araújo (2018, p. 84-85, grifo nosso) considerou que o termo pode ser entendido por um tripé conceitual apoiado em três dimensões, a saber: física, cognitiva e sociocultural.

> a) O primeiro conceito de informação na ciência da informação é mais restrito e está vinculado à sua dimensão material, **física**, sendo o fenômeno estudado a partir de uma perspectiva quantitativa e positivista.
> b) Nos anos seguintes, tomou corpo um conceito um pouco mais amplo voltado para a dimensão **cognitiva**, sendo informação algo associado à interação entre dados (aquilo que existe materialmente) e conhecimento (aquilo que está na mente dos sujeitos), e seu estudo relacionado à identificação de significados, interpretações;
> c) Por fim, as tendências contemporâneas implicam um grau maior de complexidade e abstração, com a inserção da informação no escopo da ação humana e no âmbito de contextos **socioculturais** concretos.

Visando o resgate conceitual do termo “informação”, e baseado nas dimensões apontadas acima, Capurro (2003) apresentou, nas discussões do Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB), três dimensões que, epistemologicamente, contribuíram para mapear dimensões que

alicerçam a cientificidade da CI. Essas dimensões, que contextualizam desde a teoria matemática aos estudos fenomenológicos, são pautadas nas propriedades advindas dos estudos realizados à época, e trazem consigo um aporte empírico-epistêmico, tal como se observa no Quadro 3.

Quadro 3 **-** Síntese das dimensões de Rafael Capurro.

DIMENSÃO	AUTORES	SÍNTESE
Física	Claude Shannon Michael Buckland Norbert Wiener Warren Weaver	"Tem suas raízes e seu significado nas atividades clássicas de bibliotecários e documentalistas". Esta dimensão está relacionada à Teoria da Informação e à cibernética. "Postula que existe algo, um objeto físico, que um emissor transmite para um receptor". "Essa teoria tomada como modelo de ciência da informação implica uma analogia entre o transporte físico de um sinal e a transmissão de uma mensagem cujos aspectos semânticos e pragmáticos intimamente relacionados ao uso diário do termo informação ..."Ainda soma-se às essas perspectivas o conceito de 'informação como coisa', "ou seja, algo tangível como documentos e livros ou, mais geralmente, qualquer tipo de objeto que possa ter um valor informativo".
Cognitivo	BC Brookes Henri Lafontaine Karl Popper Nicholas Belkin Paul Otlet Pertti Vakkari Peter Ingwersen	"A documentação e a ciência da informação aparentemente têm a ver primeiro com os portadores físicos do conhecimento, mas, na realidade, seu objetivo é a recuperação da própria informação, isto é, do conteúdo desses portadores". "Essa teoria parte da premissa de que a busca por informações tem sua origem em uma necessidade que surge quando existe o estado cognitivo anômalo mencionado, no qual o conhecimento disponível para resolver um problema não é suficiente."

Social	Bernd Frohmann Birger Hjørland Hanne Albrechtsen Ludwig Wittgenstein Martin Heidegger Michel Foucault	Esta dimensão reflete o problema dos limites da dimensão cognitiva ao considerar informações como algo separado do usuário localizado em um mundo metafísico ou para ver o usuário, se não exclusivamente, se em primeiro lugar, como sujeito cognitivo, deixando de lado o condicionamento social e material da existência humana. Nesta dimensão a construção dos processos informativos deve considerar a constituição social das necessidades informacionais percebidas na dimensão cognitiva.

Fonte: O autor (2022) baseado em Capurro (2003).

Considerando a descrição das dimensões apresentadas no Quadro 3, o escopo conceitual defendido em Araújo (2018) e polos capurrianos, percebe-se que o conceito de “informação”, para além da CI, abrange outras áreas do conhecimento, como a Física e a Biologia. Tais preceitos corroboram com Capurro e Hjorland (2007, p. 164) pois eles dão ênfase à importante contribuição de outras áreas na formação epistemológica da CI, e, ponderando essas afirmações, os autores conceituam “informação” em três grandes áreas: nas Ciências Naturais, Humanas e Sociais.

Nas Ciências Naturais “[...] informação significa, no nível humano, o conceito, não o processo de pensamento em si. Para que um conceito seja informação, duas condições são necessárias, a saber: deve ser uma entidade linguística e deve ser unívoco” (CAPURRO; HJORLAND, 2007, p. 164). A fim de tornar mais evidente essas condições, será abordada a Teoria do Conceito, proposta por Dahlberg, que reflete sobre esses dois aspectos.

Nas Ciências Humanas e Sociais, Capurro e Hjorland (2007, p. 172-173) afirmam que “[...] a informação é definida como um predicado que qualifica diálogos padronizados em que as locuções linguísticas são invariantes em relação a quem fala, a quem ouve e à forma” ainda, considerando os campos de estudos que constituem essas ciências, é dito que “[...] informação é um conceito-chave em sociologia, ciência política e na economia da assim chamada sociedade da informação”.

Para um conceito geral do termo “informação”, ainda pautando-se de Capurro e Hjorland (2007, p. 187), é válido considerá-lo como “[...] qualquer coisa que é de importância na resposta a uma questão”. Nesse sentido, os autores (p. 187) nos

levam a afirmar que “[...] qualquer coisa pode ser informação”. Contribui-se ainda para este pensamento, afirmar que

> Na prática, contudo, informação deve ser definida em relação às necessidades dos grupos-alvos servidos pelos especialistas em informação, não de modo universal ou individualista, mas, em vez disso, de modo coletivo ou particular. Informação é o que pode responder questões importantes relacionadas às atividades do grupo-alvo. A geração, coleta, organização, interpretação, armazenamento, recuperação, disseminação e transformação da informação deve, portanto, ser baseada em visões/teorias sobre os problemas, questões e objetivos que a informação deverá satisfazer. (CAPURRO; HJORLAND, 2007, p. 187-188)

Na CI, o termo “informação” pode estar vinculado a diferentes extratos do conhecimento, a depender de vários fatores que alicerçam os estudos desenvolvidos. Esses fatores podem estar ligados às correntes teóricas, aos autores dominantes do objeto investigado, às propriedades informacionais do estudo, aos mecanismos metodológicos, às redes de autorias e à colaboração científica, aos diversos grupos de pesquisas das mais diversas áreas do conhecimento etc.

Historicamente, conforme observado na Introdução desta tese, a CI tem uma forte relação com os fundamentos biblioteconômicos e herdou, em consequência disso, a noção de “conhecimento” defendida pela Biblioteconomia (SILVA, 2017, p. 53).

Não obstante, as definições do conceito de “informação” fazem fronteiras com o conceito de “conhecimento”, e é, neste sentido, que o autor desta tese resgata as considerações de Le Coadic (2004), ao perceber ‘informação” como um elemento de sentido. Para tanto, conforme North e Presser (2011, p. 15-16):

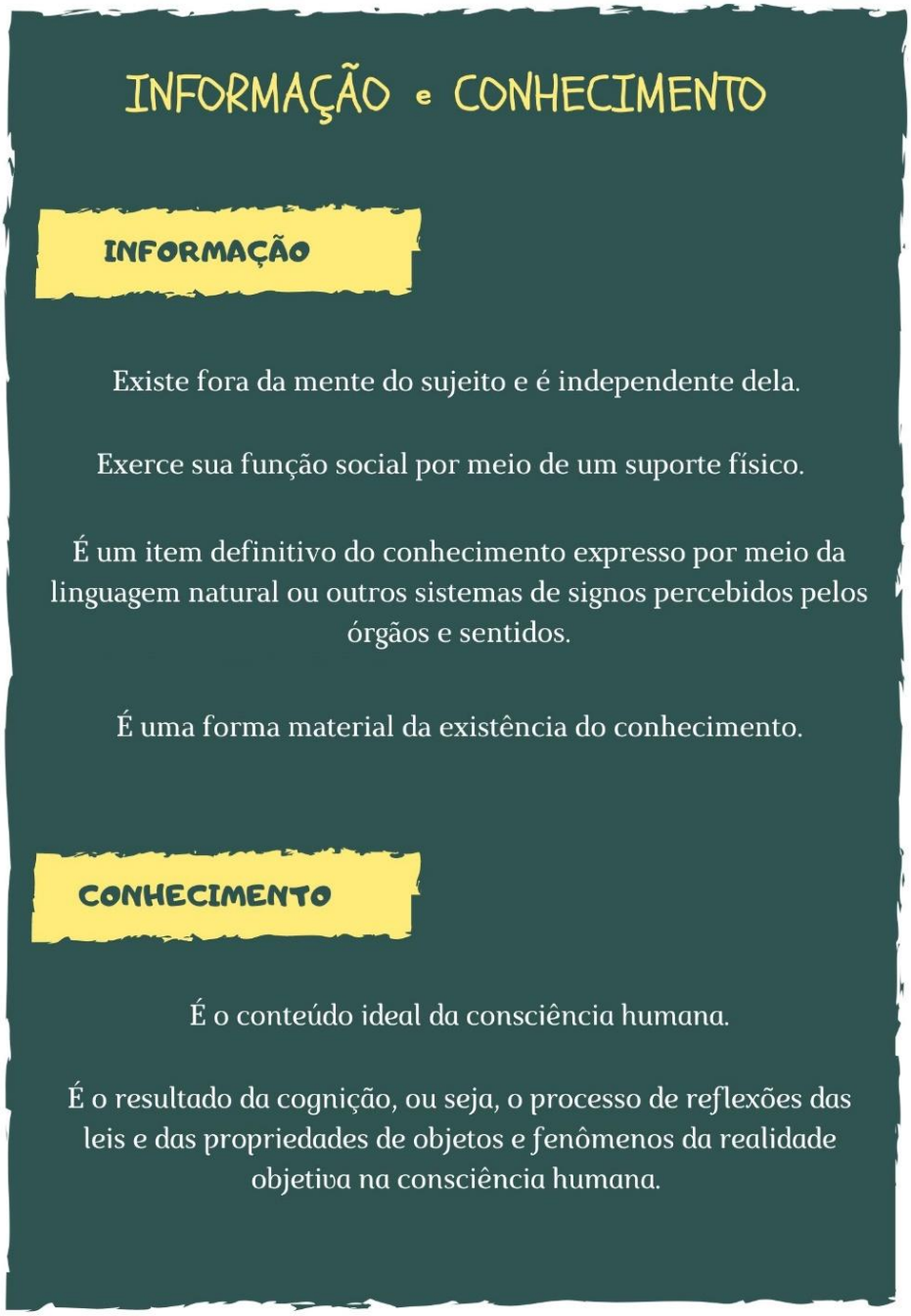
> [...] a informação só tem sentido a partir do momento em que altera a estrutura cognitiva do usuário, o receptor da informação: “informação é assim o veículo do conhecimento”, e, é por meio da informação que, quando interpretada e dotada de sentidos, se constrói o conhecimento.

As discussões, até aqui suscitadas, tornam cognoscível, portanto, que o conceito “informação” é fronteiriço ao de “conhecimento”. Assim, por vezes, “informação” e “conhecimento” são considerados sinônimos. Brascher e Café (2008, p. 3) salientam que “[...] questões relacionadas à ambuiguidades conceituais, que ocorrem nas diversas áreas do conhecimento, devem ser refletidas com base na sua legitimação pela comunidade”.

Na tentativa de relacionar, ainda que diferenciando os conceitos de “informação” e “conhecimento”, Brascher e Café (2008, p. 4) sintetizam

características advindas por Fogl (1979), apresentando as considerações ilustradas na Figura 4.

Figura 4 - Características dos conceitos de informação e de conhecimento



Fonte: O autor (2022) com base em Brascher e Café (2008, p. 4)

Findam-se, portanto, por meio desta figura conceitual, as breves discussões que limitam e relacionam os conceitos e características dos termos “informação” e

“conhecimento” desta subseção. Entende-se ainda que, por mais bem relacionadas que estejam as propriedades inerentes a eles, as nuances percebidas nos permitem diferenciá-los e, tal como, diferenciar as condições pragmáticas de suas atuações. Para esta tese essas condições estão ligadas ao conceito de Organização e Representação do Conhecimento.

2.2 Organização e Representação: da informação ao conhecimento

A busca pela definição dos termos “informação e “conhecimento” é um desafio constante para profissionais e pesquisadores da CI, tais como: bibliotecários, documentalistas, arquivistas, museólogos e gestores da informação, pois os problemas não estão concentrados, mas espaçados nos mais diversos eixos de estudo, se apoiando nos aspectos interdisciplinares da CI.

Da mesma maneira que “informação” e “conhecimento” são tratados como conceitos equivalentes, verifica-se o predomínio dessa similitude na aplicação dos termos Organização da Informação, Representação da Informação, Organização do Conhecimento e Representação do Conhecimento, cujos conceitos ocupam parte consideravelmente importante no sistema referencial da CI (BRASCHER; CAFÉ, 2008).

Concebe-se, portanto, que, se existem características distintas, quanto ao conceito de “informação” e “conhecimento”, vide figura 4, qualquer ação que tome por objeto um dos dois termos deve, igualmente, ser tratada diferentemente da que trate do outro objeto; a Organização da Informação e a Organização do Conhecimento, por exemplo, não devem ser encaradas do mesmo modo.

Assim, corrobora-se com Lima (2020) ao considerar que o termo *Organização e Representação do Conhecimento e da Informação* tem uma dispersão terminológica e são utilizados ora como complementares e ora como distintos.

Logo, considerando o exposto, esta subseção traça um cenário conceitual sobre as ações predominantes nesta pesquisa, partindo da Organização e Representação da Informação à Organização e Representação do Conhecimento.

2.3 Organização e Representação da Informação

Cabe destacar que, de acordo com as discussões introdutórias desta pesquisa, as três perspectivas epistemológicas, ilustradas na Figura I, servem de base para a sustentação da Ciência da Informação, a qual tem como gênese as atividades orientadas à organização e à representação. Na história da Documentação, que contribui para a institucionalização da CI, teve, como primeira preocupação, a resolução dos problemas documentais, e, posteriormente, passou a ocupar-se com o conjunto de sistemas, fundamentos e técnicas de organização (SOUZA, 2015).

Destaca-se, ainda, que a Documentação, ao vivenciar, historicamente dois períodos, pré e pós Segunda Guerra, primeiramente estava direcionada à organização dos registros bibliográficos e ao acesso a esses registros a partir das tecnologias disponível na época, como os microfilmes; contudo, no segundo momento, a Documentação centralizou suas atividades na análise e no tratamento do conteúdo, acentuando o uso das tecnologias voltadas à recuperação da informação (SOUZA, 2015).

A recuperação da informação é, portanto, uma função da Organização da Informação, uma vez que essa serve como instrumento para produção e disseminação do conhecimento - enquanto aquela fornece métodos e estratégias para lidar com o volume de informações (PANDO, 2018).

Considera-se, portanto, que o conceito de Organização da Informação é uma das atividades de maior relevância na Ciência da Informação, por atuar como uma ação mediadora entre as informações e o atendimento às necessidades informacionais (PANDO, 2018), possibilitando o acesso ao conhecimento contido na informação (BRASCHER; CAFÉ, 2008). Para que essa atividade seja executada,

> [...] é preciso realizar a descrição física e de conteúdo dos objetos informacionais. A descrição de conteúdo tem por objeto o primeiro dos três elementos da informação propostos por Fogl – o conhecimento. A descrição física, por sua vez, direciona-se ao terceiro elemento – o suporte da informação. O segundo elemento – a linguagem – permeia os dois tipos de descrição. (BRASCHER; CAFÉ, 2008, p. 5)

Logo, estabelece-se, como contribuição conceitual à Organização da Informação, percebê-la como um processo, um processo descritivo, cujo esse é

resultado de duas operações[6]: *descrição física*[7], que é a identificação das informações em um determinado documento: autoria, título, suporte, data de publicação, o qual também conceitua-se como metadados; e *descrição dos conteúdos*[8] dos objetos informacionais, que é a identificação dos assuntos abordados e as formas de representar esse conteúdo através dos termos, ou seja, visa caracterizar o documento do ponto de vista do seu conteúdo (BRASCHER; CAFÉ, 2008; BRAZ, 2020; DIAS; NAVES, 2013). A representação da informação é, por assim dizer, o produto das atividades de Organização da Informação.

> [...] representação da informação consiste em uma prática de enunciação das propriedades de um objeto informacional (unidade de informação organizável) ou das relações desses objetos com outros que o identificam, através das operações que visam sua descrição. Nesse entendimento, os objetos informacionais podem ser representados levando em consideração tanto as operações que contemplem a sua descrição física (também denominadas de tratamento descritivo ou representação descritiva), como também a descrição do seu conteúdo temático (tratamento temático ou representação temática) para fins de organização, acesso e uso. (ALBUQUERQUE; GAUDÊNCIO; SANTOS, 2019, p. 17)

A representação da informação é uma tarefa indispensável para organização, recuperação, acesso e uso dos objetos informacionais e, em virtude disso, são criados processos, produtos e instrumentos que alicerçam tal atividade. Tão logo, a representação da informação deve estar presente nos sistemas de organização e representação da informação, independentemente do grau de formalismo neles encontrado e do tipo de contextos que eles representem (ALBUQUERQUE; GAUDÊNCIO; SANTOS, 2019; BRAZ, 2020).

Quando relacionada ao processo de busca e recuperação da informação, a representação da informação é composta pelos seguintes elementos: documentos ;

---

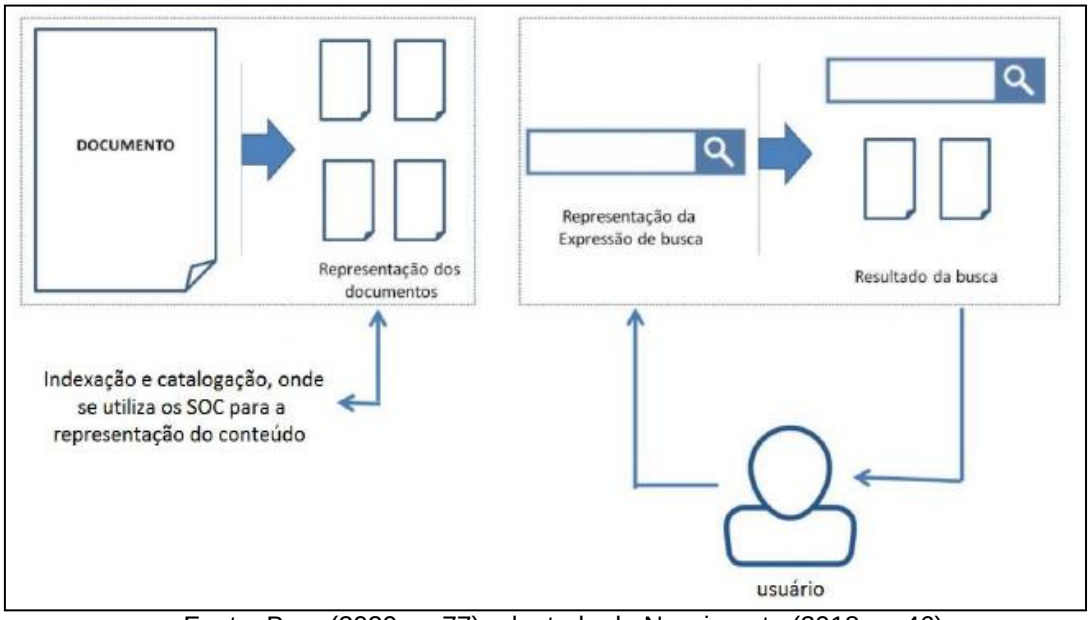
[6] Essas operações também podem ser conceituadas como tratamento descritivo, quando relacionada à descrição física e tratamento temático, quando relacionada à descrição do conteúdo ou descrição temática (DIAS; NAVES, 2012, p. 6; ALBUQUERQUE; GAUDÊNCIO; SANTOS, 2019, p. 17).

[7] "O processo de elaborar a descrição física de um documento compreende, em primeiro lugar, um exame do documento com o objetivo de identificar certos elementos nele constantes e que servirão para identificá-lo. Se o documento traz, de forma clara, todos os elementos necessários a esse tipo de identificação, a tarefa será, de modo geral, relativamente fácil. Entretanto, há casos bastante complexos que podem demandar extensas pesquisas em fontes de variadas naturezas (documental, institucional etc.) para se estabelecerem as informações precisas" (DIAS; NAVES, 2012, p. 9).

[8] "Na descrição temática, faz-se outro tipo de exame, embora também chamado de leitura técnica, apenas que, agora, com um objetivo diferente: identificar o assunto ou assuntos do documento. Esse exame constitui a fase inicial da chamada análise de assunto, primeira grande etapa da descrição temática. As etapas seguintes incluem processos como o de extração e seleção de conceitos, até que se chegue a determinar o assunto do documento" (DIAS; NAVES, 2012, p. 10).

termos de indexados; estruturas gráficas; e dados de catalogação (MOTA; KOBASHI, 2019, p. 184). A fim de ilustrar essa relação, apresenta-se a Figura 5.

Figura 5 - Representação da Informação no processo de busca e de recuperação da informação



Fonte: Braz (2020, p. 77) adaptado de Nascimento (2018, p. 46).

Por meio da Figura 5 percebe-se que é possível a visualização dos componentes dos dois processos da representação e recuperação da informação. Neste sentido, Braz (2020, p. 78) explica que

> De acordo com a figura, é possível visualizar os componentes dos dois processos da representação: enquanto os itens são descritos conforme suas tipologias e os dados acerca de sua identificação através da catalogação e da indexação, com a utilização de Sistemas de Organização do Conhecimento (SOC) para representação do conteúdo, o resultado desse processo descritivo são os catálogos ou outro sistema que tenha a mesma finalidade de armazenar a representação de itens, tais como bases de dados e portais. Cada termo de indexação ou informação sobre a obra são considerados como pontos de acesso ao documento, ou seja, as formas possíveis de chegar até ela. Os pontos de acesso são utilizados pelo usuário, que formula sua expressão de busca, obtendo como resultado os documentos que deseja ou as informações para sua localização.

Destarte, conclui-se que “[...] no contexto da OI e da RI, temos como objeto os registros de informação. Estamos, portanto, no mundo dos objetos físicos, distinto do mundo da cognição, ou das ideias, cuja unidade elementar é o conceito” (BRASCHER; CAFÉ, 2008, p. 5).

Nesse sentido, pode-se afirmar que, enquanto a informação está ligada ao mundo dos objetos físicos, ao mundo concreto e puramente identificável, o conhecimento está ligado ao mundo da cognição.

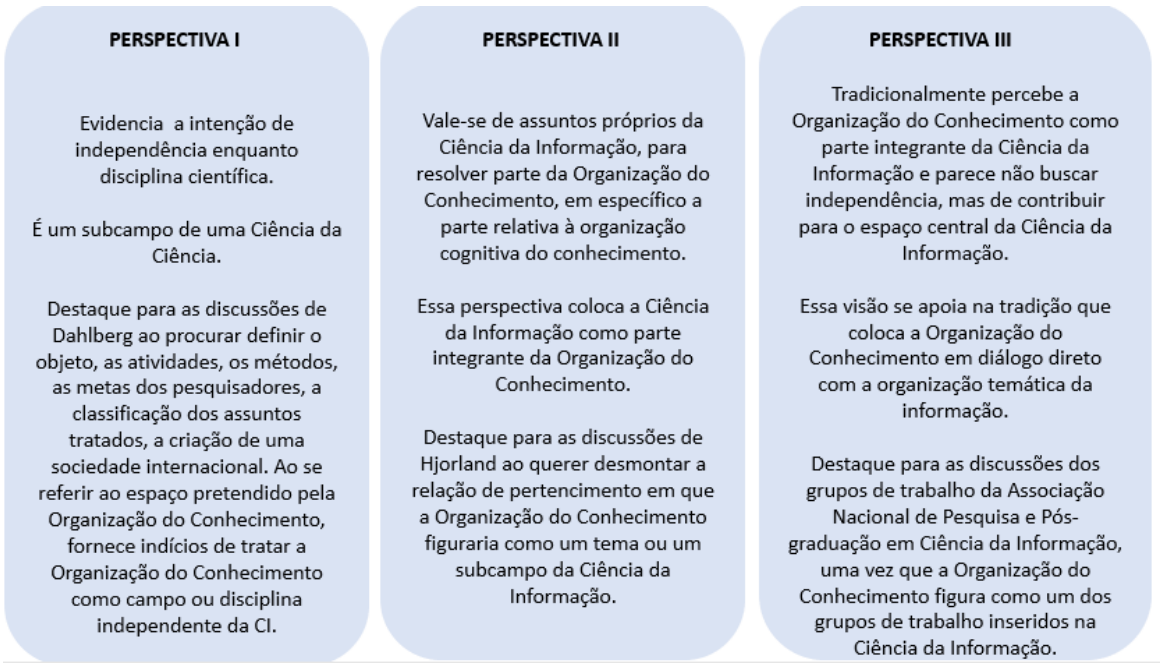
Tem-se, portanto, o uso dos termos Organização e Representação da Informação e Organização e Representação do Conhecimento, podendo ser utilizados sob vários aspectos. Possivelmente, a ideia de diferenciá-los está na concepção de que cada termo possa ser visto como processo à parte: o de organização da informação, cujo produto é sua devida representação e o de organização do conhecimento, legitimando a representação do conhecimento (BRASCHER; CAFÉ, 2008; PANDO, 2018). As Figura 9 e 10, apresentam uma perspectiva conceitual desses processos.

Com o intuito de aproximar e identificar os aportes teóricos, que constituem e diferenciam esses dois processos de organização, optou-se por tratá-los separadamente. Aqui, nesta subseção, trabalharam-se as características inerentes à Organização e à Representação da Informação. Na subseção a seguir, estão abordadas as características da Organização e da Representação do Conhecimento.

2.4 Organização e Representação do Conhecimento

Em perspectivas históricas, a OC ora é vista como uma área independente da CI, ora se relaciona e busca contribuir para os fundamentos e a aplicabilidade da CI, ora pertence à CI, não como uma área distinta, mas valendo-se de assuntos próprios da CI (SALES, 2015). Ilustra-se na Figura 6 uma melhor compreensão quanto a essas perspectivas.

Figura 6 - Perspectivas da Organização do Conhecimento e sua relação com a Ciência da Informação



.

Fonte: O autor (2022) com base em Sales (2015, p. 77-78).

Quando Sales (2015) identifica Dahlberg (Figura 7) como destaque para as discussões que orientam a Organização do Conhecimento como uma ciência, (perspectiva I), se deve ao fato do *escrito Knowledge Organization: A New Science?*, publicado na revista *Knowledge Organization* em 2006. A autora se apoia nos critérios de Alwin Diemer sobre “o que faz uma ciência”. Em arremate, Dahlberg (2006, p. 12) defende que “[...] um critério para existência de uma ciência reside no fato desta possuir um objeto próprio e, principalmente, uma área própria de atividade”, por assim defender a OC como ciência.

> O nome inclui uma combinação simples de conceitos, na qual o objeto e sua própria área de atividade já estão indicados, como conceitos de sujeito e predicado, ou seja, "conhecimento" no sentido de "o conhecido" e "organização" no sentido de a atividade de construir algo de acordo com um plano. Esses dois conceitos abrangem, portanto, a área objeto da organização do conhecimento. (DAHLBERG, 2006, p. 12, tradução nossa)

Figura 8 - Ingetraut Dahlberg

Fonte: Wikipédia (2020).

Figura 7 - Birger Hjorland

Fonte: Lima (2017).

Enquanto Dahlberg (2006) domina a perspectiva I, Hjorland (Figura 8) é o destaque da perspectiva II. O autor conceitua a organização do conhecimento sob dois sentidos, um mais restrito e outro mais amplo. No mais restrito, a OC trata de processos de descrição, indexação e classificação, sendo a OC um campo de estudo orientado à natureza e à qualidade desses processos de organização do conhecimento, bem como dos sistemas de organização de conhecimento; No mais amplo, a OC, para Hjorland (2008), é sobre a divisão social do trabalho mental, isto é, a organização de universidades e outras instituições de pesquisa e ensino superior, a estrutura de disciplinas e profissões, a organização social da mídia, a produção e a disseminação do conhecimento etc. (HJORLAND, 2008, p. 86).

Em abordagem nacional, conforme destacado na perspectiva III, a OC é tratada como uma área de estudos voltada às atividades de ordenação, representação e recuperação da informação. No âmbito da CI, a OC possui forte ligação com os aspectos epistemológicos da Biblioteconomia (PINHO, 2006). Vale ressaltar que a organização do conhecimento tem seus fundamentos na base teórica da classificação[9] (LIMA, 2015).

> A área de Organização do Conhecimento (OC) tem suas raízes históricas na base teórica da Classificação, sendo que, às vezes, pode ser vista como uma área de conhecimento, a qual classifica o conhecimento filosófico, que se inicia com o modelo categorial proposto por Aristóteles; e, às vezes, como classificação bibliográfica, que possui processos e instrumentos para

[9] “A classificação pode ser vista como uma área de conhecimento, a qual classifica o conhecimento filosófico, que se inicia com o modelo categorial proposto por Aristóteles; e como classificação bibliográfica que possui processos e instrumentos para representar conteúdo temático de documentos e aumentar a eficácia na recuperação de informações. As duas abordagens da classificação procuram ligar teoria e prática nos estudos e pesquisas realizadas dentro da Biblioteconomia e Ciência da Informação” (LIMA, 2015, p. 670).

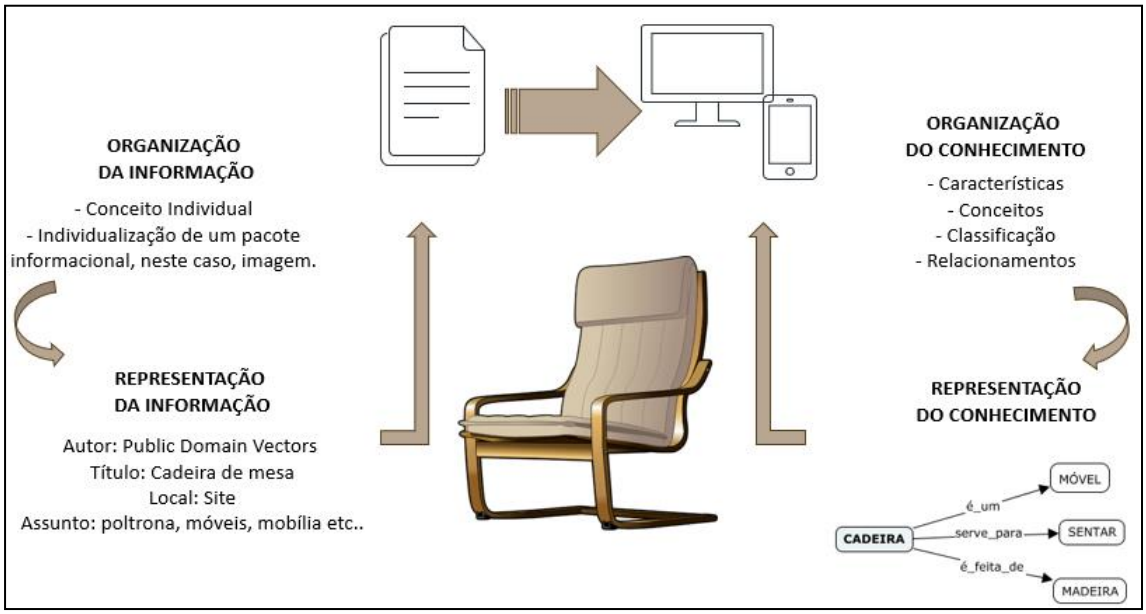
> representar conteúdo temático de documentos e aumentar a eficácia na recuperação de informações. (LIMA, 2020, p. 62)

Assim, para esta tese, na reunião dessas perspectivas epistemológicas da relação entre Organização do Conhecimento e Ciência da Informação, concebe-se que a OC se configura como um domínio de estudos que recebe espaço na CI, [...] não apenas como uma atividade pragmática, mas se amplia como área que contempla estudos teóricos e aplicados que objetivam tornar mais acessível o uso e a apropriação do conhecimento produzido (SANTOS; NEVES; SOUZA, 2019, p. 102).

A OC está vinculada ao processo de modelagem do conhecimento, o qual tem como objetivo a construção de representações do conhecimento. Esse processo tem por base a análise e a identificação das características dos conceitos, bem como suas relações (BRASCHER; CAFÉ, 2008). OC se refere, portanto, às questões conceituais que buscam compreender suas características e os relacionamentos, que podem estabelecer-se entre os conceitos, sejam eles gerais ou específicos (BRAZ, 2020).

Na Figura 9 estão ilustradas as quatro perspectivas conceituais tratadas nesta subseção, representando os conceitos de Organização e Representação da Informação e Organização e Representação do Conhecimento.

Figura 9 - Propriedades conceituais de Organização e Representação da Informação e do Conhecimento.

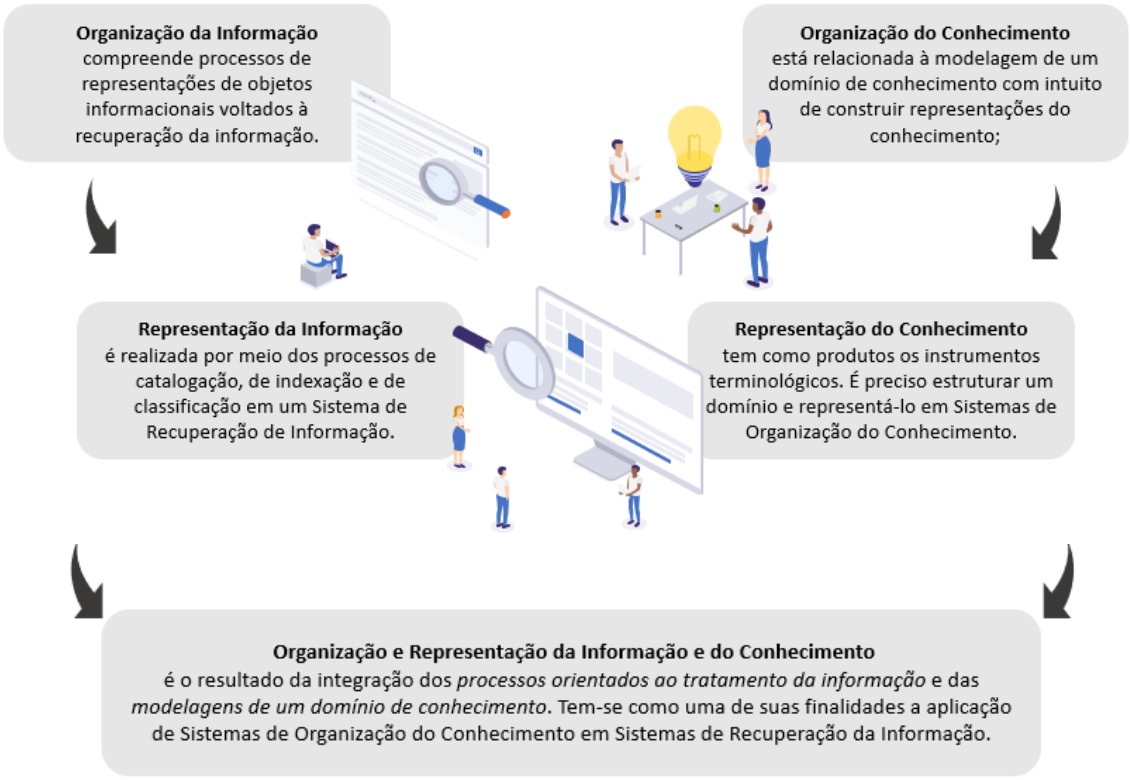


Fonte: O autor (2022) adaptado de Brascher e Café (2008, p. 7).

Ao observar a Figura 9, percebe-se que existem dois processos de organização e representação: o primeiro está orientado à característica individual do objeto informacional, para arranjá-lo sistematicamente, fazendo com que se entenda que a Organização da Informação produz a representação da informação que, por sua vez, é vista como o conjunto de atributos (conhecidos também como metadados ou pontos de acesso) que representam determinado objeto informacional, por meio dos processos de descrição física e de conteúdo; o segundo, de organização do conhecimento, está debruçado sobre a identificação dos determinantes gerais do objeto, tal como suas características, seus conceitos e suas relações. Esta, então, constitui a representação do conhecimento, vista como uma estrutura conceitual que representa modelos de mundo (BRASCHER; CAFÉ, 2008).

Diante do que fora explorado até aqui, para esta tese, considera-se como campo conceitual de *organização e representação da informação e do conhecimento* as discussões refletidas por Lima (2020). Essas reflexões estão representadas na Figura 10.

Figura 10 - Concepções conceituais de Organização e representação da informação e do conhecimento.



Fonte: O autor (2022) baseado em Lima (2020)

De acordo com as figuras vistas anteriormente, é evidente que se pode considerar que organização e representação da informação, tal como organização e representação do conhecimento, podem ser explicadas isoladamente, isto é, em suas propriedades e concepções particulares. Contudo, é necessário compreender que essas abordagens refletem, também, em uma concepção de união, de complementariedade. Assim,

> [...] pode-se dizer que essas duas abordagens se complementam, pois, enquanto a área de OC organiza o conhecimento a partir da modelagem de domínios, representando sua estrutura com base na descrição de conceitos e dos relacionamentos semânticos entre eles, na área de OI ocorre o processo de representação do conteúdo de um documento específico, o qual tem o conhecimento registrado, para que ele seja posteriormente recuperado pelo usuário em um Sistema de Recuperação da Informação (SRI). (LIMA, 2020, p. 63)

Um elemento em comum nos processos de Organização da Informação e do Conhecimento é a unidade “conceito”. Naquele, é visto como um elemento individual

do objeto que representa; neste, como a reunião de características pertencentes a um objeto. O estudo intitulado *Teoria do Conceito* (1978), originado por Ingetraut Dahlberg, pode ser utilizado como ponto de partida para uma compreensão mais consistente acerca da representação do termo “conceito”.

2.5 A Teoria do Conceito de Dahlberg

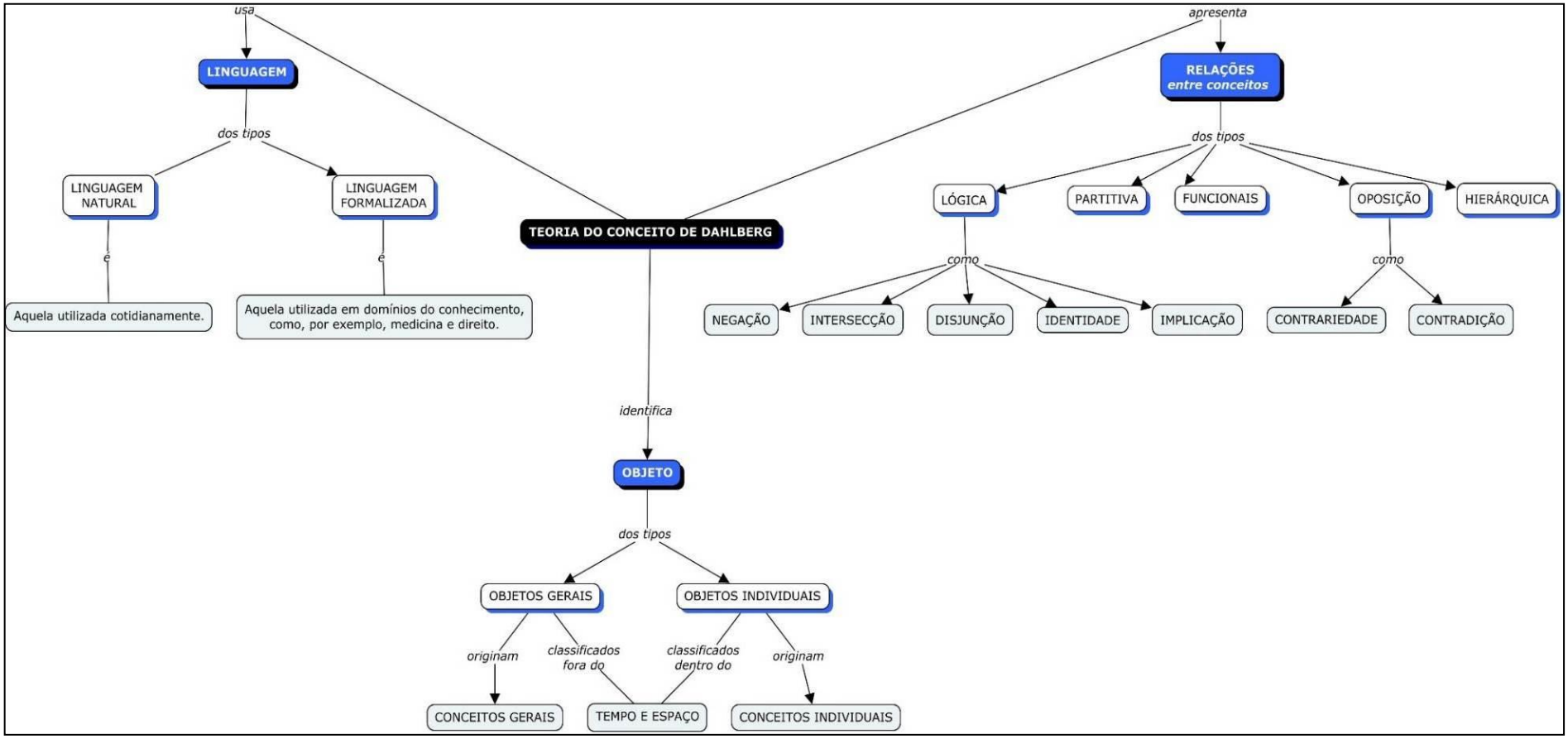
A pesquisadora alemã Ingetraut Dahlberg (1927 – 2017; figura 7) dedicou sua vida aos estudos científicos voltados à classificação e à organização do conhecimento. Sua tese, intitulada *Foundations of universal organization knowledge*, defendida em 1973, é exemplo de grande contributo às discussões que propiciam o desenvolvimento de uma fundamentação para a determinação e a compreensão dos conceitos aqui estudados (ALMEIDA; DIAS, 2019; ROCHA; CAMPOS, 2020). Através de sua pesquisa, originou-se a Teoria do Conceito, elaborada no âmbito do trabalho no Comitê de Análise Conceitual e Terminológica da Associação Internacional de Ciência Política e da Associação Internacional de Sociologia, que tratava de estudos terminológicos para as Ciências Sociais (ROCHA; CAMPOS, 2020).

A Teoria do Conceito, de Dahlberg, está relacionada aos processos de representação e recuperação da informação e aos instrumentos que auxiliam esses processos por meio de conceitos (DAVANZO; MOREIRA, 2017). Para os autores (p. 6) “[...] a clareza sobre o conceito de ‘conceito’ é fundamental para a eficácia na representação da informação e, consequentemente, para o processo de recuperação da informação”.

Dahlberg (1978), portanto, reúne propriedades para designar os atributos que compõem um conceito, tais como: **linguagens** (naturais e especializadas); **identificação do objeto** (individuais e gerais); apoio das linguagens nas **formulações de conceitos** (individuais e gerais); **características** (acidentais e essenciais) e suas **tipologias** (matéria, qualidade, tempo etc.) pertencentes aos conceitos; **relações** entre conceitos (lógicas, hierárquicas, partitivas etc.); **sentidos** de extensão e intensão do conceito; **espécies** de conceito; e, a **definição** de conceitos.

As ilustrações adiante apresentam, genericamente, uma síntese dessas propriedades conceituais em duas partes.

Figura 11 - Teoria do Conceito de Dahlberg: linguagem, objeto e relações



Fonte: O autor (2022) com base em Dahlberg (1978).

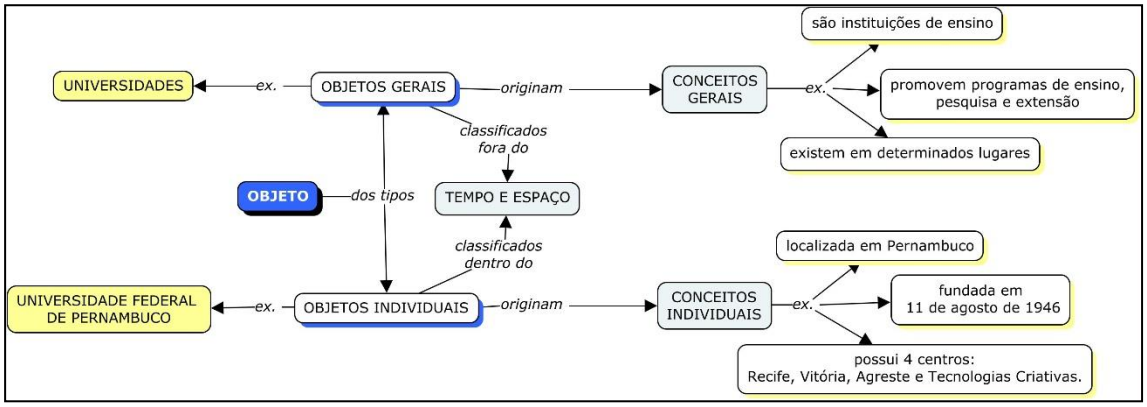
Dahlberg (1978, p. 101) fundamenta a Teoria do Conceito em determinadas propriedades, refletidas a seguir.

A **linguagem** "[...] é essencial para os humanos como seres sociais e é usada para comunicar uma visão de mundo compartilhada entre as comunidades" (THARANI, 2020, p. 221, tradução nossa).

Para efetivar uma linguagem precisa Dahlberg (1978) distingue seus aspectos em linguagens naturais que são aquelas utilizadas nas necessidades diárias e linguagens formalizadas que reúnem aspectos terminológicos orientados à um domínio do conhecimento, como a da Química, a da Medicina, e a do Direito.

Foi por meio da linguagem que o homem se tornou capaz de estabelecer relações com os **objetos** em seu entorno, estando, portanto, capaz de elaborar enunciados sobre esses objetos. Assim, quando o objeto é percebido como único, constituindo-se como unidade inconfundível, são formados os objetos individuais, destacando-se a presença das formas no tempo e no espaço. Por outro lado, existem aqueles que prescindem dessas formas, sendo eles, portanto, os objetos gerais. Tais objetos originam o que é reconhecido por "conceitos individuais", ligados aos "objetos individuais", e "conceitos gerais", ligados aos "objetos gerais" (DAHLBERG, 1978). A Figura 12 ilustra tais acepções, promovendo uma melhor compreensão do fenômeno.

Figura 12 - Teoria do Conceito de Dahlberg: objetos e conceitos



Fonte: O autor (2022) com base em Dahlberg (1978).

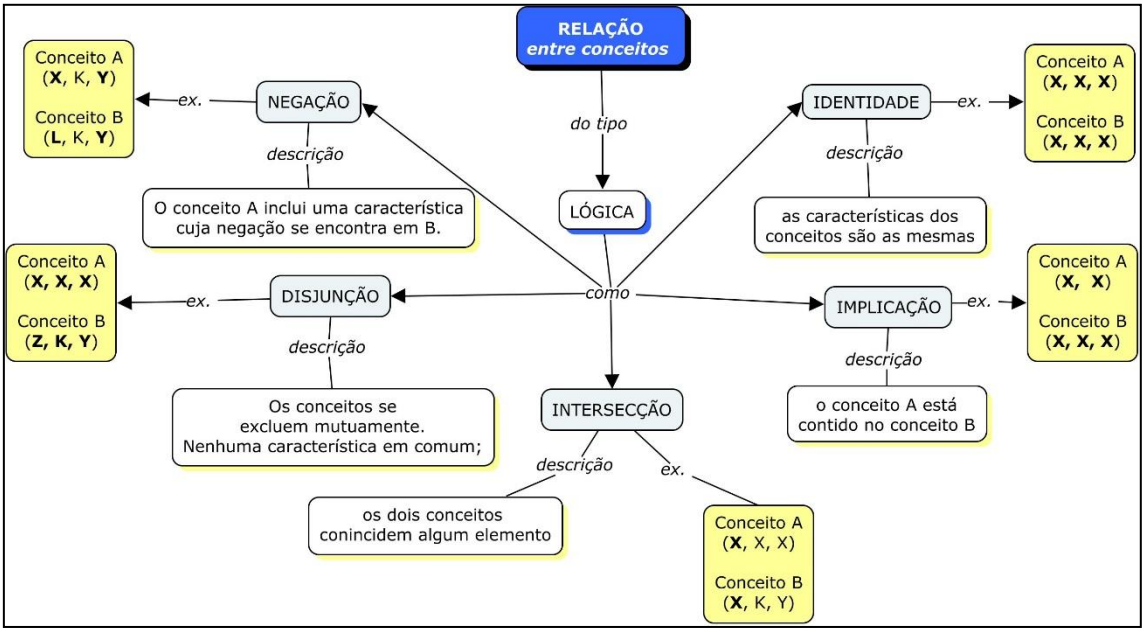
Por meio da reunião das acepções dahlberguianas, pode-se definir a formação dos conceitos como a reunião de enunciados verdadeiros sobre determinado objeto; tornando possível defini-lo como "[...] a compilação de enunciados verdadeiros sobre determinado objeto, fixada por um símbolo linguístico"

(DAHLBERG, 1978, p. 102). Esse símbolo linguístico pode ser de caráter verbal ou não-verbal, pode ser formado por sinais ou um conjunto deles, independente de palavras (DAHLBERG, 1978).

Dahlberg (1978) evidenciando que conceitos diferentes podem possuir características comuns, fala de **relações** entre conceitos. Essas podem ser: lógicas, hierárquicas, partitivas, opositivas e funcionais. Vejamos essas relações respectivamente na Figura 13.

A relação lógica é baseada nas características comuns que são logicamente possíveis. Dessa forma, com o auxílio deste tipo de relação, podem-se estabelecer comparações entre os conceitos, de modo a organizá-los em sistemas de organização do conhecimento.

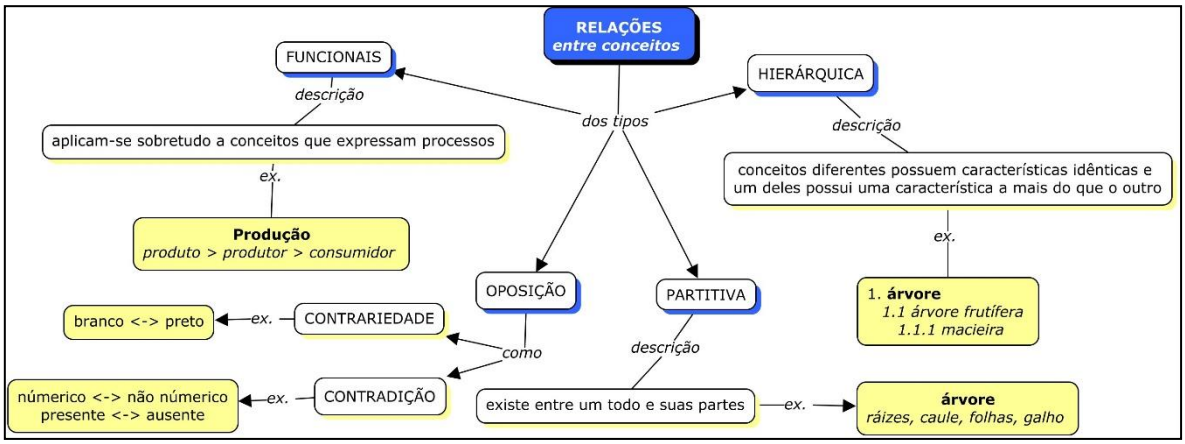
Figura 13 - Teoria do Conceito de Dahlberg: relação lógica entre conceitos



Fonte: O autor (2022) com base em Dahlberg (1978, p. 104).

Baseada em relações lógicas, Dahlberg (1978) aponta outros tipos de relacionamentos entre conceitos, a saber:

Figura 14 - Teoria do Conceito de Dahlberg: relação hierárquica, partitiva, oposição e funcional entre conceitos



Fonte: O autor (2022) com base em Dahlberg (1978, p. 104-105).

A relação hierárquica, por sua vez, é caracterizada pela ordenação e pela subordinação entre conceitos diferentes, afirmando que, embora os conceitos pareçam possuir características idênticas, sempre haverá a probabilidade de que um deles possua algo a mais em relação ao seu semelhante, isto também chama-se de relação de gênero e espécie, onde um conceito é mais genérico (superior) e o outro é mais específico (inferior). Assim, “[...] se falamos de macieira temos como conceito mais amplo ou superior o conceito de árvore frutífera e mais genérico ainda o conceito de árvore” (DAHLBERG, 1978, p. 104).

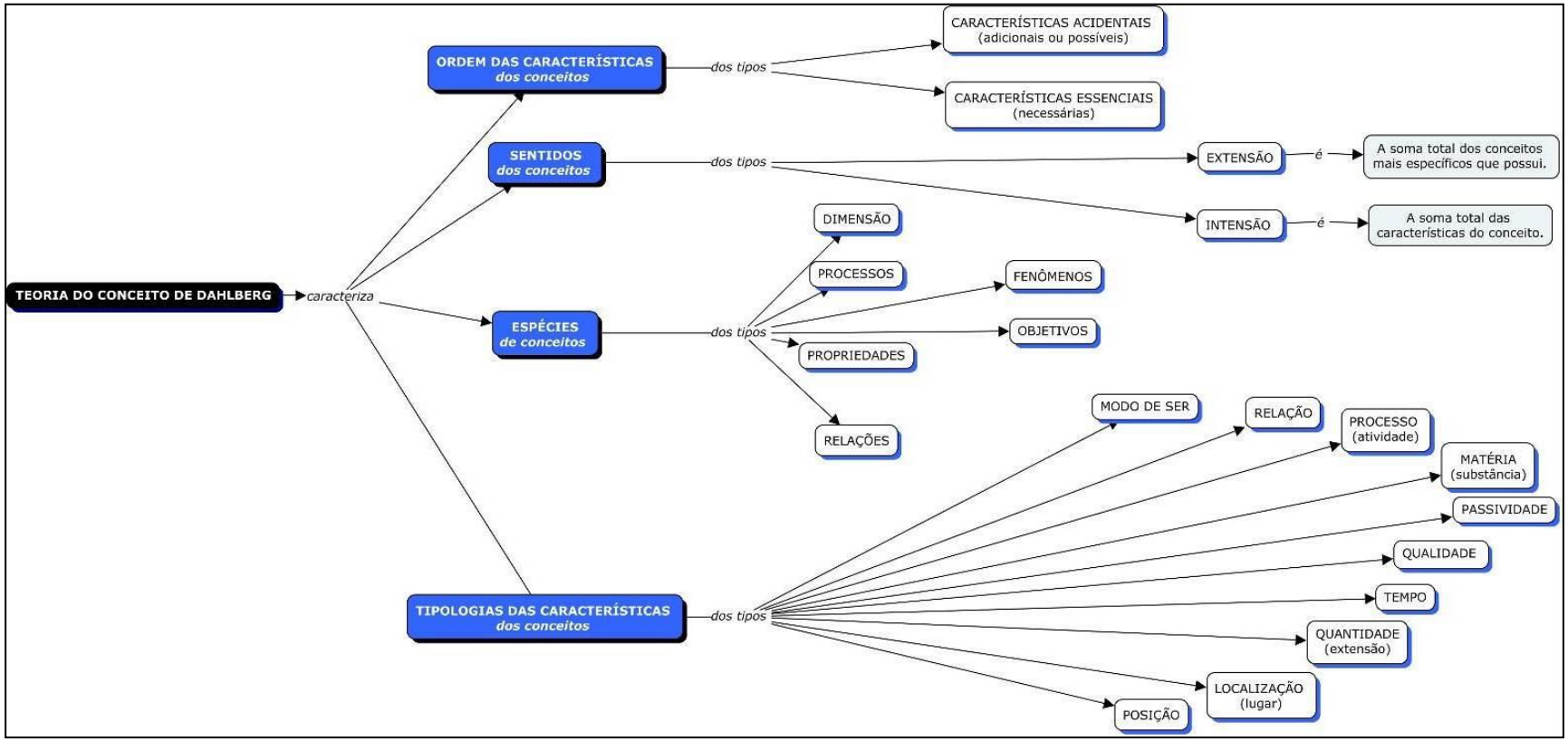
A relação partitiva é caracterizada pela razão de partição entre um objeto e os elementos que os constituem, na ilustração tomou-se como exemplo o objeto *árvore*, sendo *raízes, caule, folhas e galho*, constituintes.

As relações lógicas e partitivas aplicam-se, principalmente, aos conceitos que expressam objetos. Já as relações de oposição, aplicam-se, majoritariamente, aos conceitos que indicam propriedades, e, para este tipo de relação, são consideradas razões de contradição e de contrariedade.

A relação funcional, por sua vez, aplica-se a conceitos que expressam processos. “Pode-se conhecer o caráter semântico de tais relações tendo por base as chamadas valências semânticas dos verbos, dando atenção aos verbos e respectivos complementos” (DAHLBERG, 1978, p. 105).

Enquanto a Figura 11 representou a Teoria do Conceito quanto a linguagem, o objeto e as relações, a Figura 15, a seguir, busca representar as características dos conceitos quanto às ordens, às espécies, aos sentidos e às tipologias, observe:

Figura 15 - Teoria do Conceito de Dahlberg: ordem, espécie, sentido e tipologia de conceitos

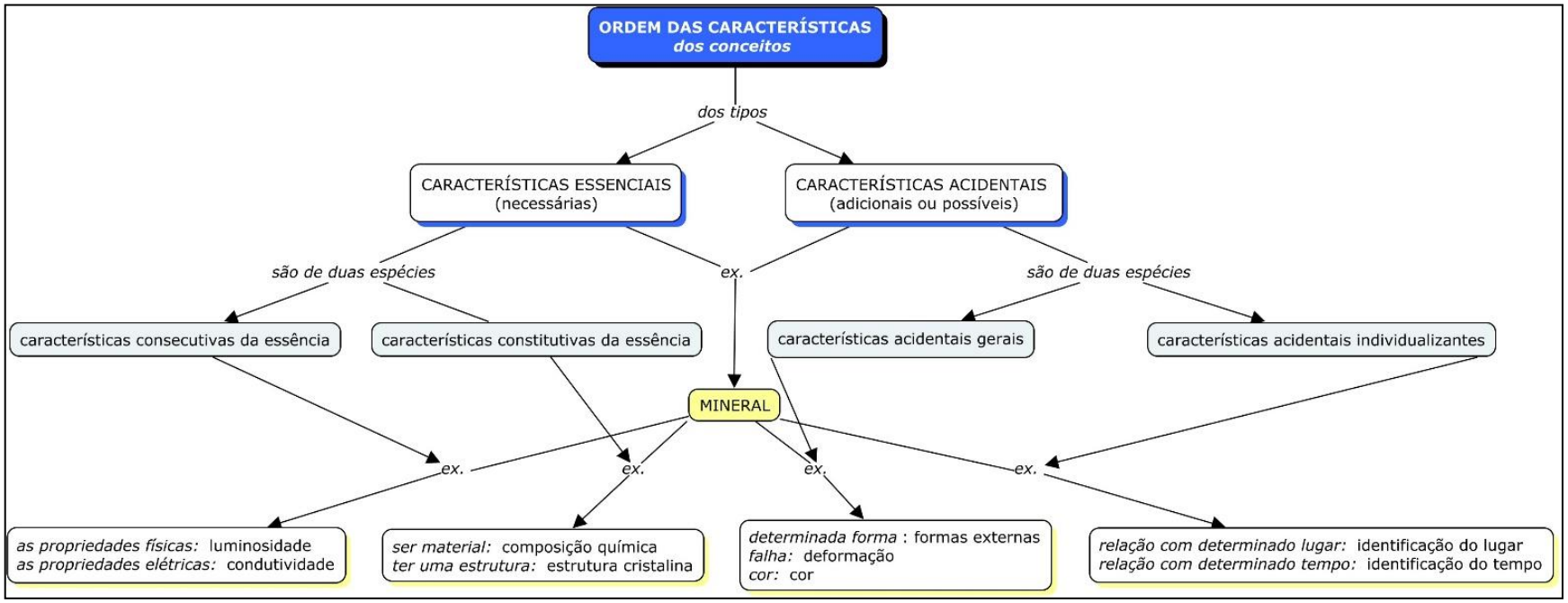


Fonte: O autor (2022) com base em Dahlberg (1978).

Como já expressado, em Dahlberg (1978, p. 102) “[...] convém repetir que é formulando enunciados sobre os atributos necessários ou possíveis dos objetos que se obtém as características dos respectivos conceitos”. A autora considera esse processo de decomposição como “análise do conceito”, “[...] mas só é possível proceder a essa decomposição do conceito coletando-se os enunciados verdadeiros que sobre determinado objeto se podem formular.” Ressalta-se, portanto, que (p. 103) “[...] se o conceito ainda não tem nome é possível formulá-lo pela síntese das características descobertas”.

A Teoria do Conceito descreve a ordem das características dos conceitos, vide Figura 16, distinguindo-a em duas razões, a primeira trata-se das características essenciais, que são determinadas pelas finalidades e pela aplicação; a segunda diz respeito às características acidentais, determinadas pela eficiência e por outros valores que possam ser, a ela, agregados.

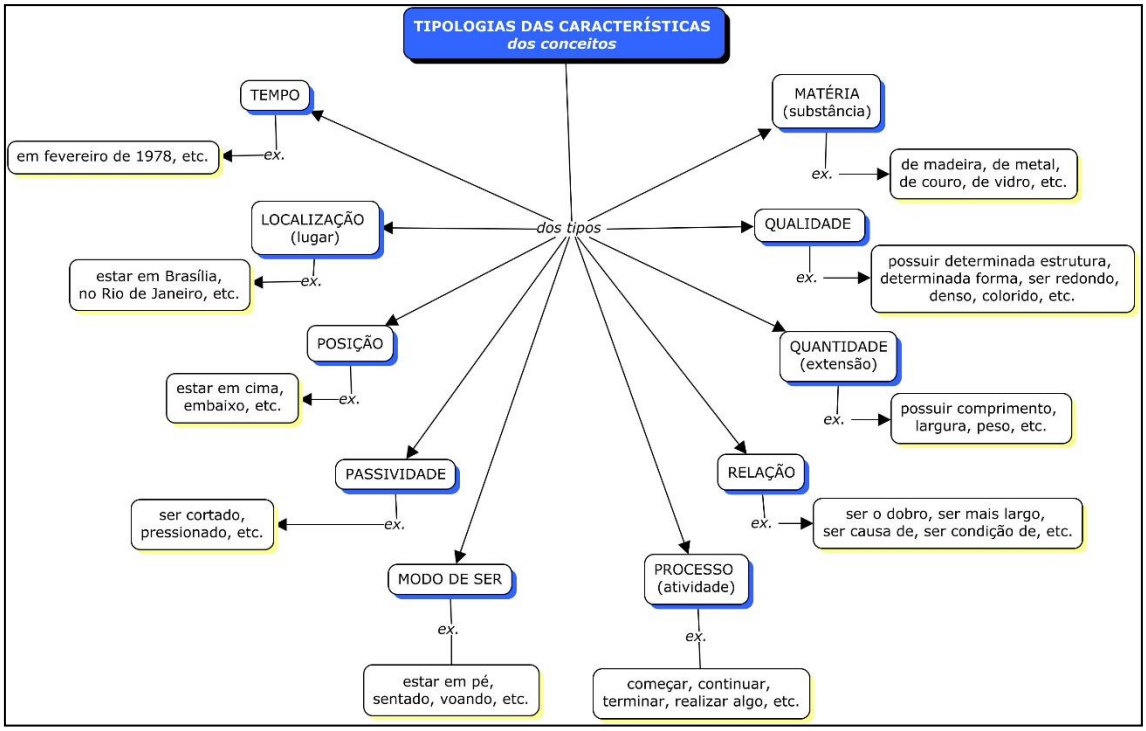
Figura 16 - Teoria do Conceito de Dahlberg: ordem das características dos conceitos



Fonte: O autor (2022) com base em Dahlberg (1978, p. 103).

Quanto às tipologias das características dos conceitos, “[...] podemos distinguir entre características simples e características complexas. São consideradas simples as que se referem a uma única propriedade” e “[...] complexas são as características que dizem respeito a mais de uma característica”, assim, “[...] em ambos os casos se trata de um material combinado com um processo resultando numa propriedade” (DAHLBERG, 1978, p. 103). A autora toma as categorias aristotélicas como base para exemplificar as possibilidades de listagem de todas as características possíveis, conforme Figura 17.

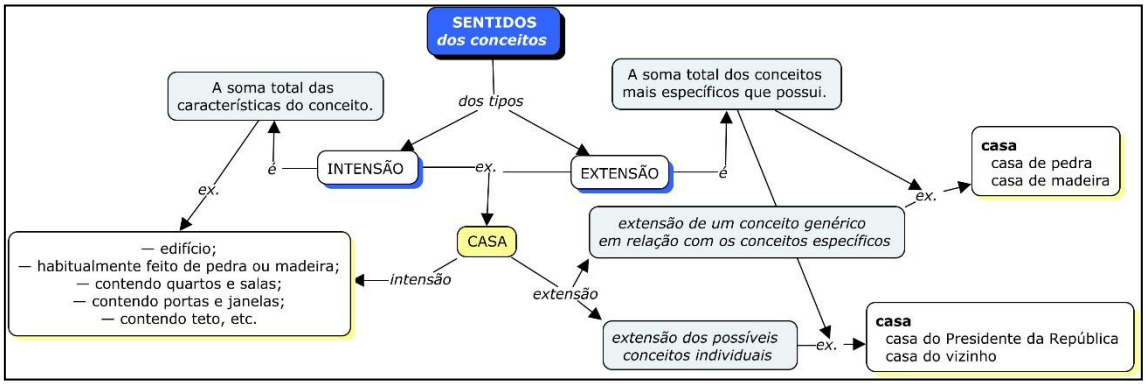
Figura 17 - Teoria do Conceito de Dahlberg: tipologias das características dos conceitos



Fonte: O autor (2022) com base em Dahlberg (1978, p. 103).

Uma outra reflexão feita pela Teoria do Conceito diz respeito aos elementos de sentido, conhecidos por intensão e extensão do conceito. A, intensão do conceito é a soma total das características, bem como dos respectivos conceitos genéricos (ou gerais) e das diferenças específicas; a extensão do conceito é a soma dos conceitos mais específicos (DAHLBERG, 1978). Figura-se o elemento de sentido do conceito:

Figura 18 - Teoria do Conceito de Dahlberg: sentido de intensão e extensão do conceito.

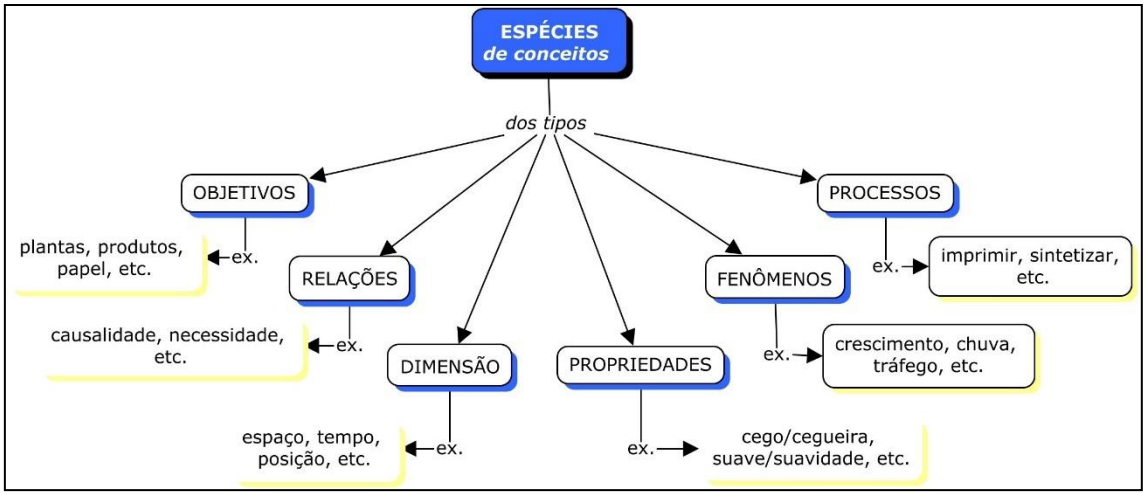


Fonte: O autor (2022) com base em Dahlberg (1978, p. 105).

Vale ressaltar que na “representação da intensão do conceito numa definição nem todos os conceitos genéricos necessitam ser mencionados” e que “[...] é suficiente mencionar o mais próximo já que este necessariamente contém os demais” (DAHLBERG, 1978, p. 105).

Nesse sentido, considerando o último elemento representado nesta tese, as espécies de conceito estão figuradas na ilustração seguinte:

Figura 19 - Teoria do Conceito de Dahlberg: espécies de conceito



Fonte: O autor (2022) com base em Dahlberg (1978, p. 105).

“Uma categorização formal dos conceitos que tem importância na formação dos sistemas e na combinação dos mesmos conceitos pode ser a seguinte: Objetos — Fenômenos — Processo — Propriedades — Relações” (DAHLBERG, 1978, p. 105). A autora relaciona este tipo de categorização às classificações facetadas, por utilizar este processo para realizar uma categorização conceitual.

Por meio das reflexões orientadas pela Teoria do Conceito, percebe-se que todos os elementos, ilustrados nesta subseção, exercem um papel fundamental na determinação do conceito, tido como elemento indispensável para se decidir o que deve ser representado e recuperado por sistemas de informação. Portanto, faz-se evidente a importância dessa teoria para a Organização e Representação da Informação e do Conhecimento, pois ela fundamenta os aspectos conceituais e suas relações, em um sistema conceitual, e possui uma relação direta com o desenvolvimento de SOC (MELO; BRASCHER, 2014; ROCHA; CAMPOS, 2020). SOC é tema da próxima seção.

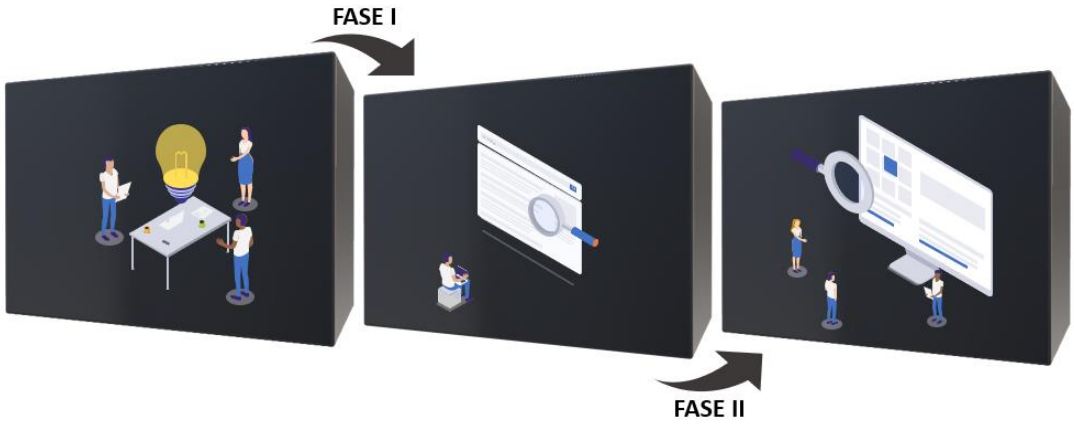
3 SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO

O termo Sistemas de Organização do Conhecimento abrange os mais diversos tipos de esquemas para organização e representação da informação, podendo, também, estender-se a outros campos da Ciência da Informação, como a Gestão da Informação e do Conhecimento. Esses sistemas incluem esquemas de classificação e categorização, com o objetivo de organizar objetos. Assim, SOC são considerados mecanismos para organização e representação da informação e do conhecimento (HJORLAND, 2015; HODGE, 2000; MAZZOCCHI, 2018).

3.1 Conceitos e características de SOC

Os SOC representam um domínio ou subdomínio de um determinado campo do conhecimento, ao mesmo tempo que apresentam uma estrutura significativa, permitindo a exploração por meio de termos, associações e atributos. Entende-se que os SOC estão representados por duas fases, a primeira *trata-se de materializar o conhecimento em recursos de informação*; na segunda, *os recursos de informação são materializados nas estruturas propostas*, ou seja, em SOC (SUÁREZ SÁNCHEZ, 2017, p. 5). Essas duas fases estão representadas na figura abaixo.

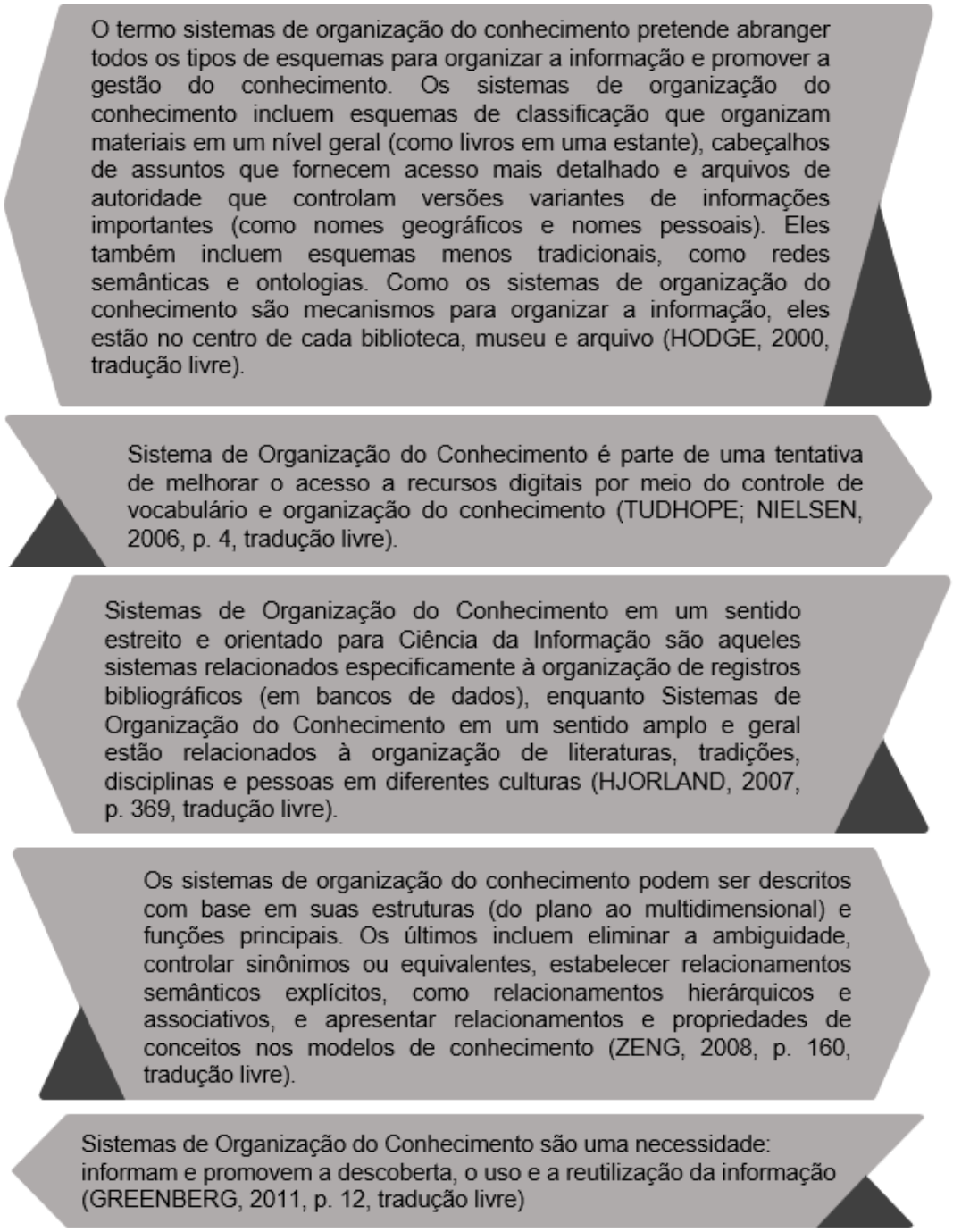
Figura 20 - As duas fases de SOC

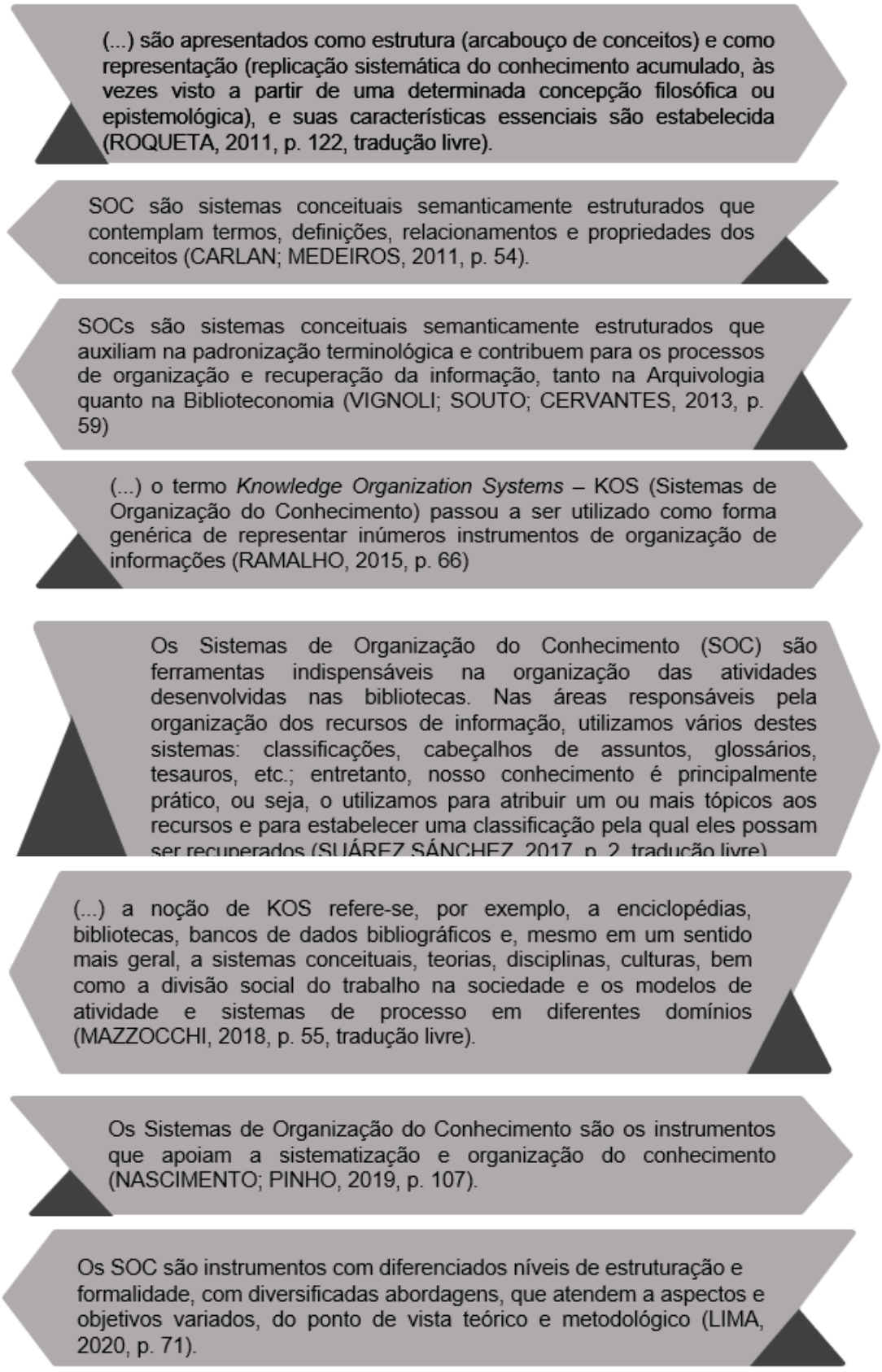


Fonte: O autor (2022) com base em Suárez Sánchez (2017).

No sentido de estreitar as discussões em torno da definição de SOC, foram reunidos estudos que apresentaram definições em seu escopo. O esquema abaixo, portanto, apresenta uma perspectiva conceitual sobre SOC.

Figura 21 - Perspectivas conceituais de SOC.



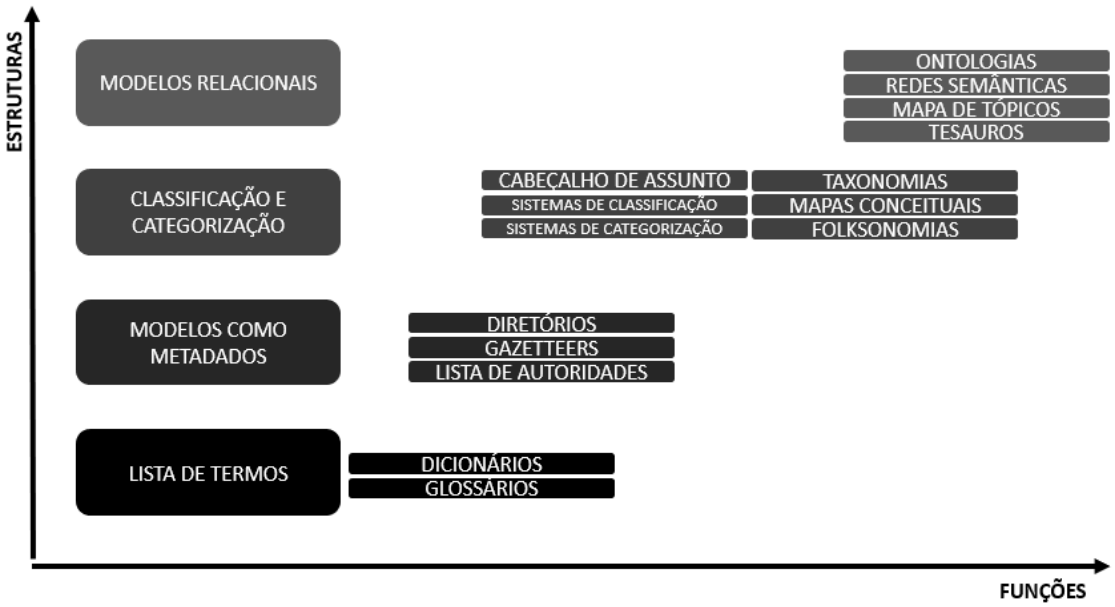


Fonte: O autor (2022) com base nas pesquisas referenciadas.

Apesar de apresentar, nesta tese, concepções acerca dos sistemas de classificação, mapas conceituais e folksonomias, e aprofundar discussões em torno dos conceitos e processos aperfeiçoados em razão do aprimoramento tecnológico digital dos sistemas taxonomia, tesauro e ontologia, pode-se, ainda, citar outros exemplos de SOC, como: glossários, dicionários, lista de autoridades, diretórios, *gazetteers*, cabeçalho de assunto, sistemas de categorização, redes semânticas e mapa de tópicos (CARLAN, 2010; HJORLAND, 2015; HODGE, 2000; MAZZOCCHI, 2018; MOREIRO GONZÁLEZ 2011; THARANI, 2020; ZENG, 2008).

OS SOC podem ser classificados de acordo com quatro grupos principais, das estruturas mais simples, como o grupo de *lista de termos*, às mais complexas, como o grupo dos *modelos relacionais*, e de acordo com suas funções (HODGE, 2000; ZENG, 2008). Observe a figura a seguir.

Figura 22 - Exemplos de SOC quanto a estruturas e funções



Fonte: O autor (2022) com base em (CARLAN, 2010; HJORLAND, 2015; HODGE, 2000; MAZZOCCHI, 2018; MOREIRO GONZÁLEZ 2011; ZENG, 2008).

Quanto ao conceito dos sistemas (grupos) observados[10]:

I) **Lista de Termos**: os *dicionários* são “[...] listas alfabéticas de termos e duas definições que fornecem sentidos variantes para cada termo”; os *glossários*

[10] Os SOC não apresentados nos grupos caracterizados estão conceituados em suas respectivas subseções desta tese, como sistemas de classificação, folksonomia, mapas conceituais, além de taxonomia, tesauro e ontologia.

são “[...] listas alfabéticas de termos, gerealmente com definições”. (ZENG, 2008, p. 161, tradução nossa);

II) **Modelos como Metadados**: os *diretórios* são “[...] listas de nomes e suas informações de contato associadas” (ZENG, 2008, p. 161, tradução nossa); *gazetteers* são “[...] coleções organizadas de informações sobre itens geográficos”; *listas de autoridade* são “[...] listas de termos empregados para controlar os nomes das variantes de um item (ou o valor de domínio de um campo específico), e onde um termo é identificado como o preferido”. (MAZZOCCHI, 2018, p. 61, tradução nossa);

III) **Classificação e Categorização**: *cabeçalhos de assuntos* são “[...] vocabulários de termos controlados, que representam o assunto dos itens em uma coleção e icluem regras para combinar esses termos em cabeçalhos compostos”; os sistemas de categorização são “[...] esquemas de agrupamento fracamente montados” (MAZZOCCHI, 2018, p. 61, tradução nossa);

IV) **Modelos relacionais**: *redes semânticas* são “[...] sistemas onde termos/conceitos são modelados como em uma rede de tipos de relacionamento variável” (MAZZOCCHI, 2018, p. 61, tradução nossa); *mapa de tópicos* é “[...] um sistema padronizado para descrever estruturas de informação e melhorar a localização dos conceitos e objetos de informação na web semântica” (MOREIRO GONZÁLEZ, 2011, p. 112).

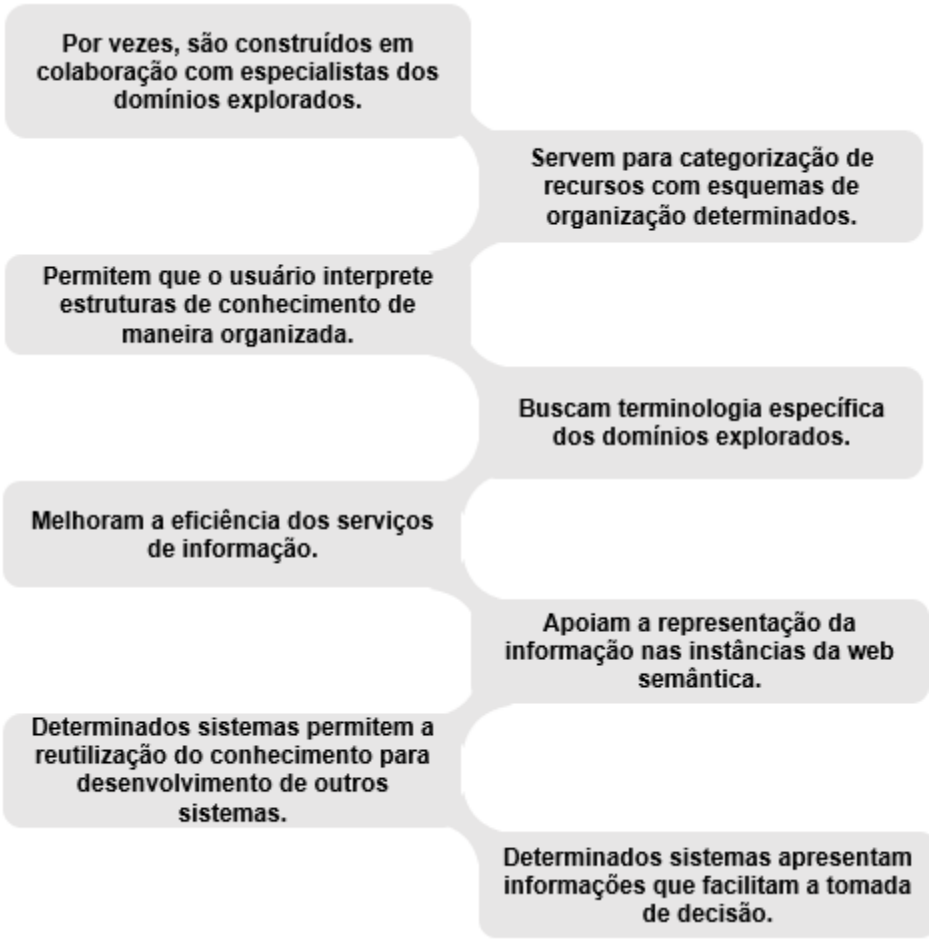
Nesse sentido, foi possível representá-los de forma a percebê-los desde estruturas mais simples às estruturas mais complexas, tanto quanto suas funções, tais como: eliminação de ambiguidades, controle de sinônimos, estabelecimento de relações e apresentação de propriedades (ZENG, 2008). Somam-se às funções atribuídas aos SOC: permitem controlar as diferentes variantes da denominação de uma entidade ou um domínio específico; destinam-se a organização de coleções por classes, de acordo com assuntos tratados (MOREIRO GONZÁLEZ, 2011).

Nota-se, no entanto, que existem funcionalidades em um determinado SOC que podem não estar presentes em outros, isso porque as finalidades de desenvolvimento e aplicação desses sistemas devem ser especificadas. Essa diferença deve-se ao fato de que, “[...] embora alguns atributos possam ser úteis em

uma variedade de situações, não há recurso ou função que, em princípio, seja adequado para todos os contextos possíveis” (MAZZOCCHI, 2018, p. 70).

Uma vez contextualizados os aspectos que caracterizam os SOC, tais como conceitos, estruturas e funções, abaixo ilustram-se as contribuições que esses sistemas são capazes de apresentar.

Figura 23 - Contribuições dos SOC.



Fonte: O autor (2022) com base em (LIMA, 2020; MOREIRO GONZÁLEZ 2011; NASCIMENTO; PINHO, 2018).

Assim, cabe-se considerar que, para o desenvolvimento desses sistemas, “[...] estão listados itens com estruturas e conteúdos diversos, originados em contextos distintos de uso para propósitos diversos e expressão de diferentes abordagens teóricas e metodológicas” (MAZZOCCHI, 2018, p. 57). Assim, é percebido que existem estruturas distintas para desenvolvimento de cada um dos SOC citados nesta seção.

A construção de SOC é vista como uma “[...] tarefa árdua, tanto intelectualmente como nos tempos necessários ao seu desenvolvimento” (MOREIRO GONZÁLEZ; SÁNCHEZ CUADRADO; MORATO LARA, 2012, p. 47). Em uma concepção aplicada, esses sistemas “[...] são representações esquemáticas do conhecimento, úteis para ordenar informações alojadas em entidades ou recursos” (SUÁREZ SÁNCHEZ, 2017, p. 6).

A Teoria do Conceito, abordada na seção anterior, contextualiza os fundamentos que alicerçam o desenvolvimento e o funcionamento de SOC. Esses sistemas são dotados do conjunto de propriedades percebidas por Dahlberg (1978), a saber: linguagem como elemento mediador; identificação do objeto conceitual como elemento de percepção primeira; formulação dos conceitos, bem como suas devidas características; e tipologias.

Os elementos que mais motivam o desenvolvimento de SOC são as propriedades de relacionamento, que, por sua vez, integram os conceitos representados em um SOC e os relacionamentos entre os elementos desses sistemas.

Cabe ressaltar que os elementos-chave de investigação da tese são: tesauro, taxonomia e ontologia. Todavia, apontam-se, nesta pesquisa, os demais tipos de SOC trabalhados na CI, como as Classificações, os Mapas Conceituais e a Folksonomia. Não se pretende, contudo, esgotar as discussões desses últimos SOC, mas construir uma percepção de existência e de legitimidade para cada um deles, como contributo para a Ciência da Informação e para a Organização da Informação e do Conhecimento.

Classificações é o primeiro SOC a ser estudado.

3.2 Classificações

Inicialmente, é válido destacar, mais uma vez, que o estudo deste SOC, nesta pesquisa, se limita a abordar breves contributos que contextualizem a importância das Classificações Bibliográficas para a CI e para o campo de Organização da Informação e do Conhecimento. Exploraram-se dois textos que norteiam esta subseção, o livro *Introdução à Teoria da Classificação* (1983), de Maria Antonieta Requião Piedade e a tese de Marcos Luiz Cavalcanti de Miranda, intitulada

Organização e Representação do Conhecimento: fundamentos teóricos-metodológicos na busca e recuperação da informação em ambientes virtuais (2005), para desenvolver um breve itinerário sobre as Classificações.

Este itinerário inicia-se conceituando: “o que é classificar?”. O termo classificar tem origem latina, *classis,* que designava os grupos em que o povo romano se dividia. Foi dicionarizado, em 1733, por Zedler no *Universal Lexicon*, que combinava as palavras latinas *classis* e *facere*, para apresentar uma divisão de apelações de Direito Civil e, posteriormente, passou a ser visto como um termo empregado para divisão das ciências (PIEDADE, 1983, p. 15).

Piedade (1983, p. 16) define “classificar” como a divisão

> [...] em grupos ou classes, segundo as diferenças e semelhanças. É dispor os conceitos, segundo suas semelhanças e diferenças, em certo número de grupos metodicamente distribuídos [...] “é um processo mental habitual ao homem, pois vivemos automaticamente classificando coisas e ideias, a fim de compreender e conhecer.

“Classificação” é vista como uma das atividades inerentes ao homem desde sua existência. É um ordenamento de sentidos, objetos e fenômenos que busca, para além da complexa ação de classificar, a representação das relações sociais e científicas existentes. Corrobora com este conceito o pensamento de Miranda (2005, p. 107) ao refletir que “[...] os seres humanos estão conectados por uma rede complexa de relações sociais. Todo indivíduo na sociedade está *linkado* com diversos outros por vínculos sociais”.

> Como sabemos, o processo de classificar é inerente aos seres humanos desde cedo, em contato com nosso mundo através da comparação entre objetos novos ou da experiência com aqueles com que estamos familiarizados, identificando padrões e categorizando o que é novo em nosso quadro de referência existente. (MIRANDA, 2005, p. 118)

Miranda (2005) sintetiza um contexto científico centralizado em Kaula (1985), em que produziu um estudo verticalizado sobre classificação e

> Explana o conceito de classificação como entendido e exposto pelos filósofos o que inclui os Predicados de Aristóteles, a Árvore de Porfírio, a Classificação Baconiana e os esquemas escolásticos para a classificação do conhecimento como dado por John Locke, Auguste Comte, Herbert Spencer etc. Demarca as categorias dos diferentes esquemas de classificação, as quais incluem os esquemas modernos e que têm sido utilizados para a organização do conhecimento registrados em livros e armazenados em bibliotecas. Discute a falácia da classificação de livros e as razões para se adotar o termo classificação bibliográfica. Aponta a falácia da classificação bibliográfica. Esclarece em detalhes o background teórico dos diferentes esquemas de classificação partindo da CDD até a

> RIC (Rider International Classification). Analisa as diferentes abordagens no campo da classificação e suas implicações. Fornece as características do universo de assuntos e a abordagem estática dos classificacionistas anteriores a Ranganathan. Comenta a atual teoria dinâmica da classificação. Traça a análise conceitual na classficiação e as visões sobre a aplicação da teoria de Ranganthan para a CDU referindo-se ao design da ASCOM (Analitic-Sintetic Classification of Medicine). Miranda (2005, p. 120)

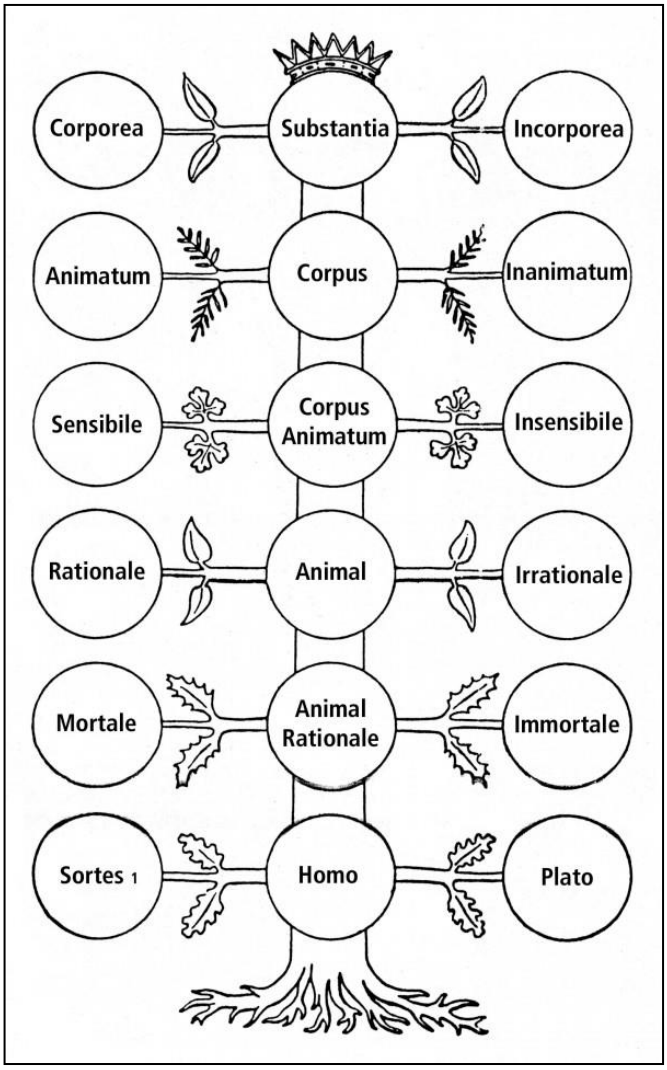
Piedade (1983), a fim de esclarecer o que são os predicados de Aristóteles[11], diz que o filósofo nomeou de categorias ou predicáveis as classes gerais, em que se situam, ordenadamente, às ideias que temos das coisas e que constituem os dez gêneros supremos. Esses gêneros são: substância; qualidade; quantidade; relação; duração; lugar; ação; paixão ou sofrimento; maneira de ser; e posição.

Aristóteles buscou agrupar os dez gêneros citados e os reduziu à três: substância (é o ser que existe), modo ou acidente (aquilo que existe na substância) e relação (aquilo que liga um ser ao outro) (PIEDADE, 1983). Na concepção dessas categorias, deparamo-nos com a Árvore de Porfírio[12]. A Árvore de Porfírio pode ilustrar a concepção dessas categorias, posto que é tida como ponto de partida para os estudos da Classificação, segundo Miranda (2005).

[11] Como exemplo ilustrativo já visto nesta tese, tem-se a figura *Teoria do Conceito de Dahlberg: ordem, espécie, sentido e tipologia de conceitos* na página 41, a qual Dahlberg se baseia nos pensamentos de Aristóteles para listar as características categóricas do conceito (DALHBERG, 1978, p. 103).

[12] “Porfírio, filósofo grego do século IV, apresentou na sua obra *Isogoge* ou Introdução às Categorias princípios para regerem as subdivisões das classificações denominados predicáveis ou categoremas, que são os cinco seguintes: Gênero, Espécie, Diferença, Propriedade e Acidente” (PIEDADE, 1983, p. 23).

Figura 24 - Árvore de Porfírio



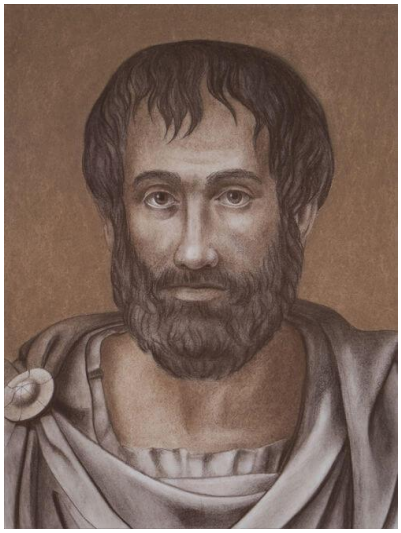
Fonte: (BORDONI; DANSA, 2012, *online*).

As Classificações, como esquemas sistemáticos, podem ser do tipo filosóficas, quando são criadas pelos filósofos e têm a finalidade de definir, esquematizar e hierarquizar o conhecimento - a este tipo de classificação está o ordenamento das ciências ou o ordenamento das coisas (PIEDADE, 1983, p. 58).

Para a Piedade (1983, p. 61), esses sistemas surgiram quando "[...] os sábios compreenderam que o Universo é um sistema harmônico, cujas partes estão dispostas em relação ao todo, que há uma hierarquia das causas e dos princípios [...] e uma relação entre as ciências que os estudam". Tendo isso em vista, Piedade

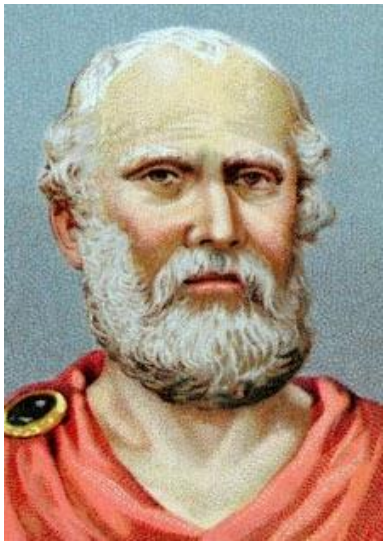
(p. 61) cita que Richardson (1930)[13] “[...] admite que a classificação teve início quando o homem começou a dar nome às coisas. O ‘ato de classificar’, diz ele, ‘fez do macaco um homem’”. O autor, baseado em Sayers, “[...] considera que o primeiro filósofo a se preocupar em classificar as ciências foi Platão (428 – 347 a. C.), que, na República, dividiu o conhecimento em: Física Ética e Lógica”, enquanto outros autores, como “Lars[14] e Dahlberg[15], preferem iniciar a história das classificações filosóficas com Aristóteles (384 – 322 a. C.)” (PIEDADE, 1983, p. 61).

Figura 25 - Aristóteles (384 – 322 a. C.)



Fonte: Porfírio (2019).

Figura 26 - Platão (428 – 347 a. C.)



Fonte: Vieira (2014).

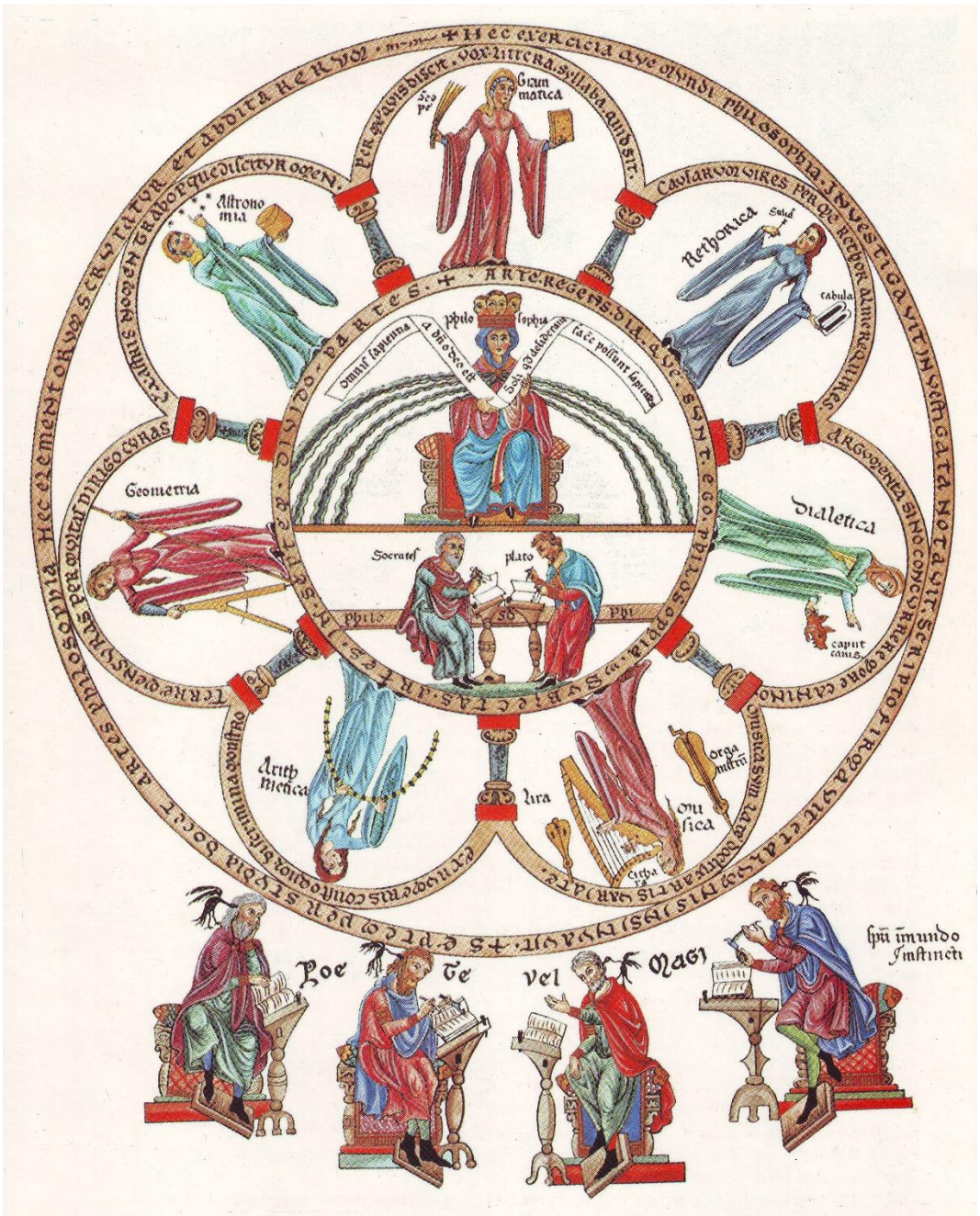
Platão classificou as ciências em Física, Ética e Lógica; Aristóteles, por sua vez, ordenou as ciências em: Teóricas, Práticas e Poéticas. Dentre estas, a primeira está voltada à constatação da verdade e tem como exemplo a Matemática, a Física e a Teologia; a segunda, isto é, a das ciências práticas, é vista como a que determina as regras que regem nossos atos, como a Moral ou a Ética, a Economia e a Política; por último, as ciências poéticas indicam os meios a serem usadas a produção de obras exteriores, tal como a Retórica, a Poética e a Dialética (PIEDADE, 1983).

[13] RICHARDSON, E. C. **Classification**: theoretical and practical. New York: Wilson, 1930.
[14] LARS, C.; JASPERS, L. **Manual de filosofia**. 7. ed. São Paulo: Melhoramentos, [s.d.].
[15] DAHLBERG, I. **Ten lectures on theory of classification**. *In*: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA E DOCUMENTAÇÃO, 8., 1975, Brasília.

Contudo, apesar do legado fornecido por Platão e Aristóteles, destaca-se a produção do filósofo romano, Matianus Capella (c. 439), que escreveu “Das Núpcias de Mercúrio com a Filologia”, presenciado pelas ciências divididas em Gramática, Dialética, Retórica, Geometria, Astronomia e Música. As quais, mais tarde, Cassiodoro (468 – 575) agrupou em duas partes[16], a primeira, das Artes ou Ciências Sermoniais (ciências das palavras), reunindo a Gramática, a Dialética e a Retórica; a segunda, Ciências Reais (ciências das coisas), contendo Geometria, Aritmética, Astronomia e Música (PIEDADE, 1983).

[16] Foi com base em Cassiodoro que as escolas da Idade Média dividiram o ensino das ciências em dois grandes grupos, o *Trivium*, contando com as *Artes ou Ciências Sermoniais;* e o *Quadrivium*, com as *Ciências Reais*. Integram o *Trivium* e o *Quadrivium* as setes disciplinas: Gramática, Dialética, Retórica, Geometria, Aritmética, Astronomia e Música, sendo as três primeiras disciplinas pertencentes ao *Trivium* e as quatro últimas disciplinas pertencentes ao *Quadrivium*, juntas essas sete disciplinas originaram o que se conhece por “As sete artes liberais” (PIEDADE, 1983).

Figura 27 - As sete artes liberais



Fonte: Wikipédia (2020b).

Os SOC bibliográfico foram fortemente influenciados pelas classificações filosóficas (BRAZ, 2020). O *Trivium* e o *Quadrivium* influenciaram a classificação bibliográfica de Konrad von Gesner (1516 – 1565), por exemplo (PIEDADE, 1983, p. 63).

Outros filósofos contribuíram significativamente para as classificações bibliográficas, como (PIEDADE, 1983, p. 63-65):

- Francis Bacon (1561 – 1626), que dividiu as ciências em Memória, Imaginação e Razão;

- Augusto Comte (1798 – 1857), que qualificou as ciências em Abstrata, quando estuda as leis gerais, independentemente dos seres concretos em que se realizam; e Concreta, quando estuda esses seres. Comte parte das Ciências Abstratas para as Ciências Concretas, dividindo-as em Matemática, Astronomia, Física, Química, Biologia, Sociologia e Moral;

- Wilhelm Maximilian Wundt (1832 – 1920) apresentando a seguinte classificação: Ciências Formais, Reais e Ciências da Natureza (Fenomenológicas, Genéticas e Sistemáticas) e Ciências do Espírito (Fenomenológicas, Genéticas e Sistemáticas. As Fenomenológicas investigam os fenômenos, nas Ciências da Natureza são: Física, Química e Fisiologia, nas Ciências do Espírito é a Psicologia. As Sistemáticas tratam dos objetos, nas Ciências da Natureza são: Mineralogia, Botânica, Zoologia etc., nas Ciências do Espírito são: Direito, Economia, Política etc., As Genéticas nas Ciências da Natureza são: Cosmologia, Geologia e História da Evolução Orgânica; e nas Ciências do Espírito é a História.

Além das categorias aristotélicas e da Árvore de Porfírio, comumente utilizadas na Ciência da Informação, para explicar a arte da Classificação, a classificação baconiana recebe grande destaque nos estudos da organização da informação e do conhecimento, ora por ser considerada “[...] esclarecedora e satisfatória para a organização do conhecimento” (MIRANDA, 2005, p. 122), ora por ser tida como “[...] a mais importante para os bibliotecários, porque influenciou várias outras classificações bibliográficas” (PIEDADE, 1983, p. 63).

Miranda (2005), considerando Kaula (1984), reúne uma série de SOC representados pela Filosofia, e para o desenvolvimento e a aplicação nas ciências específicas.

Quadro 4 - Sistemas Escolásticos de Organização do Conhecimento e Sistemas Científicos de Organização do Conhecimento

	SOC	CLASSIFICACIONISTA	ANO
Sistemas Escolásticos de Organização do Conhecimento	Classificação de Locke	John Locke	1688
	Classificação de Comte	Auguste Comte	1822
	Classificação de Coleridge	Samuel Coleridge	1826
	Classificação de Spencer	Herbert Spencer	1864
	Classificação de Bain	Alexandre Bain	1870
	Classificação de Stadler	Auguste Stadler	1896
	Classificação de Pearson	Karl Pearson	1892
	Classificação de Richardson	E. Richardson	1901
Sistemas Científicos de Organização do Conhecimento	Classificação Botânica de Lineu	Carl Lineu	1753
	Classificação Botânica de Jussieu	Antoine Jussieu	1789
	Classificação de Bentham e Hooker	Bentham e Hooker	1862
	Classificação de Engler	Engler	1900
		Carpenter	1845
		Lydekkar	1897
		Hertwig	1903
		Cuvier	1729
		Owen	1884

Fonte: O autor (2022) adaptado de Miranda (2005, p. 123 e 124, baseado em KAULA (1984)).

Piedade (1983, p. 60) considera a Classificação Bibliográfica como "[...] sistemas destinados a servir de base à organização de documentos nas estantes, em catálogos, em bibliografias etc" e que esses sistemas (p. 65) cumprem, portanto, duas finalidades: a ordenação dos documentos nas estantes ou nos arquivos e a ordenação das referências nas bibliografias ou das fichas nos catálogos.

Logo, os sistemas bibliográficos de organização do conhecimento surgem com a finalidade de organizar os livros em estantes. Os primeiros sistemas bibliográficos não possuíam notações[17]; foi com o aumento da produção bibliográfica que esses sistemas foram criados com notação (MIRANDA, 2005).

[17] Notação, aqui, é o "[...] conjunto de símbolos destinados a representar os termos da classificação, traduzindo em linguagem codificada o assunto dos documentos, e permitindo sua localização nas estantes, nos catálogos e nas tabelas de classificação" (PIEDADE, 1983, p. 38).

Quadro 5 - A notação e os Sistemas bibliográficos de organização do conhecimento

	SOC	CLASSIFICACIONISTA	ANO
Sistemas bibliográficos de organização do conhecimento sem notação	Classificação de Aldus	Aldo Manuzzi	1505
	Classificação de Gesner	Konrad Gesner	1548
	Classificação de Maunsell	Andrew Maunsell	1595
	Classificação de Naudé	Gabriel Naudé	1627
	Classificação de Garnier	Jean Garnier	1678
	Classificação de Leibnitz	Leibnitz	1718
	Classificação de Horne	Thomas Horne	1824
	Classificação de British Museum	British Museum	1836
	Classificação de Brunet	J. Brunet	1842
	Classificação de Schleiermacher	Schleiermacher	1852
	Classificação de Merlin	R. Merlim	1842
	Classificação de Palermo	Francisco Palermo	1854
	Classificação de Royal Institute	Royal Institute	1857
	Classificação de Trubner	Nicholas Trubner	1859
	Classificação de Edward	Edward Edwards	1859
	Classificação de Smith	L. Smith	1882
	Classificação de Ogle	J. Ogle	1895
	Classificação de Sonnenchein	W. Sonnenchein	1897
	Classificação de Quinn-Brown	John Quinn e James Brown	1894
Sistemas bibliográficos de organização do conhecimento com notação	Classificação de Harris	William Harris	1870
	Classificação de Schwartz	Jacob Schwartz	1871
	Classificação de Decimal de Dewey	Melvil Dewey	1876
	Classificação Expansiva de Cutter	Charles Cutter	1891
	Classificação do Sion College	Q. Milman	1886
	Classificação Decimal Expandida de Bruxelas	Instituto Internacional de Bibliografia	1905
	Classificação Racional de Perkins	F. Perkins	1882
	Classificação de Hartwig	Otto Hartwig	1888
	Classificação de Fletcher	W. Fletcher	1889
	Classificação de Bonazzi	G. Bonazzi	1890
	Classificação de Rowell	J. Rowell	1894
	Classificação Ajustável de Brown	James Brown	1898
	Classificação Científica	Usada no CILC	1901
	Classificação Internacional da Universidade de Princeton	Universidade de Princeton	1901
	Classificação da Biblioteca do Congresso	Biblioteca do Congresso dos Estados Unidos	1902
	Classificação Decimal Universal	Federação Internacional de Informação e Documentação	1905
	Classificação de Assunto de Brown	James Brown	1906

Fonte: Adaptado de Miranda (2005, p. 125, baseado em Kaula (1984)).

A busca por uma classificação universal se iniciou nos Estados Unidos, com a publicação da Classificação Decimal de Dewey (CDD) (1876) e da Classificação Decimal Universal (CDU) (1905), devida por Paul Otlet e Henri La Fontaine. A CDU originou-se a partir da CDD e "[...] é um sistema hierárquico, com base filosófica, mas no qual, graças à utilização de sinais gráficos, diz-se que surge a tentativa de classificação em facetas, cujo primeiro emprego consciente apareceu na Classificação de Dois Pontos" (PIEDADE, 1983, p. 74). Classificação de Dois Pontos teve como autor Shiyali Ramamrita Ranganathan[18].

A CDD, a CDU e a *Colon Classification* ou Classificação de Dois Pontos, bem como outras classificações bibliográficas, apresentadas no quadro a seguir, são consideradas os maiores sistemas de organização do conhecimento bibliográficos universais (MIRANDA, 2005, p. 126).

Quadro 6 - Sistemas de organização do conhecimento bibliográficos

SOC	CARACTERÍSTICAS GERAIS
Classificação Decimal de Dewey (1876)	Melville Louis Kossuth Dewey criou em 1876, o sistema de classificação bibliográfico mais utilizado e mais conhecido em todo o mundo. O sistema conta com uma ordem de classes principais que, no princípio, foi descrita com a seguinte explicação: "[...] o homem começou a pensar e a procurar uma explicação para a sua existência, e assim surgiu a Filosofia; incapaz de desvendar o mistério imaginou a existência de um ser supremo que o havia criado, surge a Religião; multiplicando-se o homem passa a viver em sociedade e vêm as Ciências Sociais; sente necessidade de se comunicar com os companheiros e cria as línguas; passa então a investigar os segredos da natureza e temos as Ciências Puras; de posse destes conhecimentos procura deles tirar proveito, aparecendo as Ciências Aplicadas; e agora, já se sentindo capaz de criar, dá origem às Artes e à Literatura; finalmente encontramos a História que conta como tudo se passou" (PIEDADE, 1983, p. 89). Hoje, a CDD conta com dez classes principais, sendo elas: Generalidades, Filosofia, Religião, Ciências Sociais, Línguas, Ciências Puras, Ciências Aplicadas, Artes, Literatura e História e Geografia.
	Charles Ammi Cutter, bibliotecário norte-americano, foi autor do segundo código de catalogação norte-americado, o Rules for a Dictonary Catalog publicado em 1876. Mas foi em 1891, enquanto trabalhava na Biblioteca do Boston Athenaeum que a Classificação Expansiva foi definitivamente

[18] Miranda (2005, p. 91) vê Ranganathan como "[...] um dos cientistas que rompeu com o paradigma conservacionista da Organização do Conhecimento, contribuindo para uma verdadeira revolução científica através de suas leis, princípios, postulados e cânones e do método analítico sintético, primeiro na Biblioteconomia e depois na Documentação e na Ciência da Informação, em torno de uma moderna teoria da classificação: a Teoria da Classificação Facetada.". Para Piedade (1983, p. 74) "Ranganathan obrigou a uma reformulação das ideias de classificação. Tem adeptos fervorosos, principalmente entre os teoristas de classificação da Inglaterra, e críticos severos, mas seu sistema de classificação é muito pouco empregado, mesmo na Índia."

Classificação Expansiva de Cutter (1891)	criada, foi assim chamada, possivelmente pela possibilidade do sistema se expandir à proporção que a biblioteca cresce. "Cutter não justificou a ordem de suas classes principais, mas alguns estudiosos, reportados por Sayers, a explicam do seguinte modo: partindo do caos, representado pelas Obras Gerais (A), presumindo-se a existência do Homem, este toma consciência da sua própria existência e desenvolve sua mente criando a Filosofia (B). interroga-se sobre sua origem e a resposta encontrada é a existência de um Deus, sugerindo a Religião (BR); interessa-se pela sua vida como individuo aparecendo a Biografia (E); examina o meio em que vive, dando origem à História (F) e à Geografia (G); estuda suas relações com os companheiros e temos as Ciências Sociais (H-K); volta sua atenção para as forças que governam a natureza, surgindo as Ciências (L-Q); preocupa-se com os meios de subsistência e surgem as Artes Úteis (R-U); e, por último, atinge o estágio de criar e aparecem as Belas Artes (W) e sua expressão máxima, a Literatura (Y)" (PIEDADE, 1983, p. 147).
Classificação da Bilioteca do Congresso (1902)	James Hanson e Charles Martel, chefes da catalogação e da classificação da Biblioteca do Congresso em meados de 1890, tomaram por guia a Classificação Expansiva de Cutter e introduziram grandes modificações, especialmente quanto à notação. O sistema da Biblioteca do Congresso é um sistema demasiadamente detalhado, mas que recorre à síntese, ao aplicar suas inúmeras tabelas auxiliares. As classes básicas são divididas no sistema por letras do alfabeto: A (General Works), B (Philosofy, Psychology, and Religion), C (Auxiliary Sciences of History), D (General and Old World History), E (History of America), F (History of the United States and British, Dutch, French, and Latin America), G (Geography, Anthropology, and Recreation), H (Social Sciences), J (Political Science), K (Law), L (Education), M (Music), N (Fine Arts), P (Languagen and Literature), Q (Science), R (Medicine), S (Agriculture), T (Technology), U (Military Science), V (Naval Science) e Z (Bibliography, Library Science, and General Information Resources).
Classificação Decimal Universal (1905)	A história da CDU tem início em 1892 com dois belgas: Paul Otlet e Henri La Fontaine, respectivamente advogado e professor. Os dois perceberam a necessidade de melhorar a organização para controlar a bilbiografia e fundaram o *Office International de Bibliographie*, em Bruxelas. O escritório teve a finalidade de organizar uma bibliografia universal, o qual chamaram de *Repertoire Bibliographique Universal*. Contudo, foi apenas em 1905 que publicaram a primeira edição do novo sistema com o título de *Manuel du Répertoire Bibliographique Universe*l. Hoje a CDU comporta dez classes principais sendo uma delas vaga, as demais seguem a ordem de Generalidades, Filosofia. Psicologia, Religião. Teologia, Ciências sociais, Matemática e Ciências Naturais. Ciências Puras, Ciências Aplicadas. Medicina. Saúde. Tecnologia. Agricultura. Cozinha e Culinária., Belas artes. Arquitetura. Música. Design. Recreação. Turismo. Esportes. Jogos., Linguagem. Língua. Linguística. Filologia. Literatura. e Geografia. Biografia. História.
	James Duff Brown, foi dos primeiros ingleses a escrever sobre Biblioteconomia e criador do único sistema de classificação geral da Inglaterra. O inglês pretendeu criar um sistema de classificação para que cada assunto tivesse uma devida localização, estando os aspectos teóricos e práticos de um mesmo tema reunidos. "A Subject Classification

Classificação de Assunto de Brown (1906)	elaborada por James Brown em 1906, tem como base a própria teoria de Brown de que toda forma de conhecimento deriva de um dos quatro grandes princípios fundamentais na ordem em que surgiram no Universo: Matéria e Força, Vida, Razão e Registro. Primeiro surgiram a Matéria e Força (Ciências Físicas), que geraram a Vida (Ciências Biológicas, Etnologia e Medicina; Biologia Econômica. Artes Domésticas), esta produziu a Razão (Filosofia e Religião. Ciências Sociais e Políticas), que deu origem aos Registros dos fatos (Língua e Literatura. Gêneros Literários. História e Geografia. Biografia)" (MIRANDA, 2005, p. 131).
Colon Classification (1933)	É considerada como o mais importante trabalho desenvolvido por Shiyali Ramamrita Ranganathan. "A Colon é empregada em bibliotecas na Índia, mas não vem sendo utilizada fora daquele país. Seu estudo, porém, é importante devido à influência que exerce sobre os estudiosos e autores de classificações, especialmente os membros do Classification Research Group da Inglaterra, que advogam as classificações facetadas" (PIEDADE, 1983, p. 193).
Classificação Bibliográfica de Bliss (1935)	"A *Bibliographic Classification* é apontada como um sistema que apresenta um dos melhores desenvolvimentos de classes encontrado em classificações bibliográficas" [...] "Bliss procurou apresentar os assuntos segundo o "consenso científico e educacional", isto é, a ordem científica e pedagógica de estudos dos assuntos, a ordem dos conceitos geralmente aceita pelos especialistas (PIEDADE, 1983, p. 182 -183). A Classificação Bibliográfica de Bliss “[...] constitui-se de 4 volumes em 3: volumes 1 e 2, letras A-K; volume 3, letras L-Z; volume 4, índice" (MIRANDA, 2005, p. 134).

Fonte: O autor (2022) com base em Piedade (1983) e Miranda (2005).

A partir deste mapeamento dos sistemas de classificação, percebe-se a necessidade de organização e representação das publicações bibliográficas existentes - processo também conhecido como controle bibliográfico, em que se busca a identificação dos itens de um acervo bibliográfico com o objetivo de promover recuperação e acesso ao conteúdo e à identificação da área do conhecimento que este possui.

Tão logo, na busca por sistemas que facilitem os processos de organização e recuperação da informação, desta vez dispensando regras notacionais ou qualquer padrão que controle quaisquer tipos de documento, imagem, som, vídeo etc., consideram-se outros tipos de sistemas de organização do conhecimento, conhecidos como Mapas Conceituais e Folksonomias. Esses sistemas estão apresentados nas subseções abaixo.

3.3 Mapas Conceituais

Conforme supracitado, os mapas conceituais não possuem regras notacionais e são pouco controlados por padrões que orientem sua construção e uso, diferentes das políticas que regem as atividades de controle bibliográfico, em que o profissional da informação necessita seguir protocolos e métodos para organização e representação da informação e do conhecimento, derivados de um item informacional.

Corroboram, neste sentido, as indicações inferidas por Moreiro González (2011, p. 109), ao que ele identifica que, nos mapas conceituais, não se aplicam controle de termos, e que os conceitos neles representados, podem tratar das mais diversas coisas, como a busca pela representação de entidades concretas ou abstratas, eventos ou coisas; e que esses mapas podem, ainda, fornecer representações de distinções entre procedimentos ou produtos.

Utilizados em demasia na área de Educação, os Mapas Conceituais são considerados ferramentas didático-pedagógicas que possibilitam a organização e a representação do conhecimento, contribuindo, significativamente, para o processo de aprendizagem. Além disso, o uso de Mapas Conceituais potencializa experiências lúdicas, fazendo com que habilidades da mente se desenvolvam; refletindo, percebendo e analisando o processo de aprendizagem por meio da relação entre conceitos e das conexões no mapa representadas (FARIAS; FARIAS, 2016).

> A utilização desses mapas pode preparar os indivíduos não apenas para absorver, mas, sobretudo, para produzir novos conhecimentos, ainda que de forma mais livre e criativa, podendo dizer até sob uma ótica desclassificatória, onde se sobressaia a preocupação com o indivíduo sistêmico, aproveitando suas experiências intelectuais ou cognitiva, tornando as pessoas conscientes do processo de construção do próprio conhecimento. (CORDOVIL; FRANCELIM, 2018, 954)

Os Mapas Conceituais surgiram por meio de um experimento com crianças, realizado pelo professor Joseph Novak, que buscou pôr em prática os aspectos da aprendizagem significativa de David Ausubel. O experimento procurou identificar o que as crianças sabiam sobre determinado assunto, para que representação prévia desse conhecimento fosse desenvolvida; então, as crianças eram ensinadas sobre o assunto e realizavam, posteriormente, um novo mapa do conhecimento recebido.

Cordovil e Francelim (2018, p. 952) interpretam o uso de mapas conceituais como ferramentas para “[...] constituir um processo de aprendizagem que é capaz de aproximar o conhecimento dos espaços de vivência com os saberes sistematizados”; e afirmam que o resultado desse processo é a construção de representações de “[...] leituras, interpretações e raciocínios desenvolvidos pelos indivíduos ao longo de sua formação”. Desse modo, concebe-se que os mapas conceituais são construções mentais refletidas sob forma de gráficos e orientadas à percepção do indivíduo sobre um contexto ao qual este fora, previamente, submetido.

Moreiro González (2011, p. 108-109) toma como base para construção desses mapas a representação de frases simples e bem estruturadas, seguindo uma sequência de “sujeito-verbo-predicado” e ainda como um relacionamento associativo entre “conceito-relação-conceito”.

Tal como outros sistemas, percebe-se, que os Mapas Conceituais estão direcionados a duas ciências: Educação, quando tratados como objeto de aprendizagem; e a Psicologia, se tidos como instrumento de desenvolvimento cognitivo. Na Ciência da Informação, os Mapas Conceituais são vistos como recursos, isto é, como documentos ou domínios que estejam representados por um conjunto de conceitos relacionados e dispostos como um grafo (MOREIRO GONZÁLEZ, 2011).

Ademais, os Mapas Conceituais também são exemplos de produto da Organização do Conhecimento, cujo campo está ligado à Ciência da Informação, e ilustra esquemas de representações “[...] partindo da menor complexidade para os de maior complexidade”; e, por analogia aos demais sistemas, os mapas “[...] possuem em suas representações controle de ambiguidade como as listas, relacionamento hierárquico como as taxonomias e relacionamentos associativos como os tesauros” (CORDOVIL; FRANCELIM, 2018, p. 952).

De acordo com as informações supramencionadas, as associações entre conceitos, nos Mapas Conceituais, não estão normalizadas, mas se referem, geralmente, ao conjunto de ideias representadas; ao objeto todo-parte refletido nos grafos; às relações dos tipos: específica, física, temporais, lógicas, hierárquicas etc. Os mapas conceituais oferecem, portanto, uma rede de relações infinitas, mais ricas

que o sistema tesauro, que pode ter algumas características inerentes aos tipos de relações feitas nesse sistema (MOREIRO GONZÁLEZ, 2011).

O uso de sistemas como os Mapas Conceituais permite o desenvolvimento de melhores práticas para organização, representação e recuperação de informação e de conhecimento, por usuários, uma vez que as relações dos conceitos, neles existentes, estão dirigidas às necessidades e expectativas de cada usuário (MOREIRO GONZÁLEZ, 2011, p. 110). Assim, os Mapas Conceituais podem ser definidos "[...] ora como um instrumento ou como uma ferramenta, ora como estratégia, ora como método ou técnica ou recurso esquemático" (RODRIGUES; CERVANTES, 2014, p. 159).

De acordo com Moreiro González (2011, p. 110), o desenvolvimento de mapas conceituais pode atender a diferentes objetivos, tais como: geração de ideias; representação de estruturas complexas (como textos longos, hipermídias, sites etc.); comunicação de ideias complexas; e para auxiliar a aprendizagem. Como exemplos de Mapas Conceituais construídos nesta tese podem ser observadas as ilustrações da subseção "Teoria do Conceito".

Os processos para construção de Mapas Conceituais, sucedem-se a partir dos seguintes passos: seleção dos conceitos a presentar no mapa; listagem dos conceitos; agrupamento dos conceitos; ordenação dos conceitos; e enlace dos conceitos, por meio de etiquetas (MOREIRO GONZÁLEZ, 2011).

Baseadas em Moreiro González (2011) e outros autores[19], Rodrigues e Cervantes (2015) apresentam, também seis processos para o desenvolvimento de mapas conceituais, sendo: identificação do tema; verificação de conceitos; ordenação dos conceitos; agrupamento e arranjo dos conceitos; determinação das relações entre os conceitos; e a revisão da estrutura do mapa conceitual.

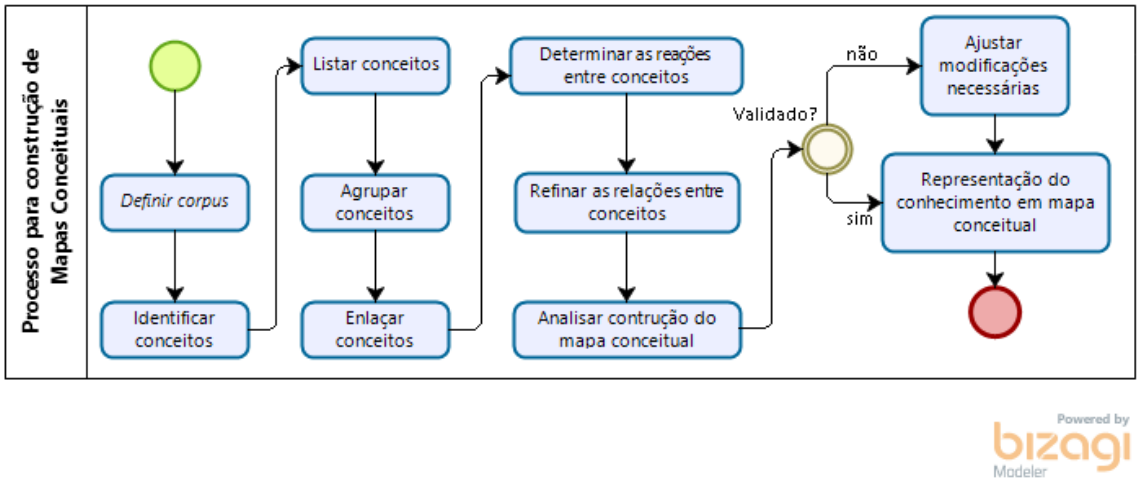
Enquanto Rodrigues e Cervantes (2015) reuniram os processos para construção de mapas conceituais em diversos outros autores, inclusive Moreiro González (2011), sobre procedimentos para o desenvolvimento de mapas, Nascimento, Braz e Pinho (2018, p. 13), quando relacionaram mapas conceituais ao domínio jurídico, reuniram os seis passos das autoras e propuseram, somado aos processos existentes, outros cinco passos:

[19] Belluzzo (2007); Lima (2004); Marques (2008); Moreira (2006); Novak (2000); Ontoria *et al*. (1999); Prats Garcia (2013).

> Análise dos documentos que se busca representar, uma análise documentária compreendendo as perspectivas linguísticas e terminológicas do domínio a ser representado (1); A seleção dos conceitos principais a fim de representá-los, considerando suas extensões (2); Refinamento das conexões entre conceitos com o intuito de garantir ligações bem estabelecidas e fundamentadas nas relações que um conceito promove ao outro (3); Análise do mapa conceitual como um todo considerando a verificação da cadeia construída, dos nós estabelecidos, das conexões conceituais bem arranjadas, percepção de inserção ou exclusão de termos ou nós (4) e, por fim, a Representação da construção do mapa conceitual (5).

Buscando uma representação processual que reúna todos os passos de construção de mapas conceituais desta seção, expõe-se, a seguir, uma tentativa de coadunar os passos em um processo linear do desenvolvimento deste sistema.

Figura 28 - Processos para construção de Mapas Conceituais



Fonte: O autor (2022) com base em Moreiro González (2011), Nascimento, Braz e Pinho (2018), Rodrigues e Cervantes (2015).

Conferindo o andamento do estudo dos SOC, seguindo as características em comum dos Mapas Conceituais, e dispensando o uso controlado por padrões que orientem sua construção e uso, o próximo sistema em estudo é a folksonomia.

3.4 Folksonomia

Diferente dos sistemas classificatórios, que dispõem métodos rígidos, para organizar o corpus e representá-lo em um sistema de recuperação da informação; e contrário, também, aos Mapas Conceituais, orientados por relações associativas e

demais outras, como as relações essenciais, temporais, lógicas etc.; a Folksonomia pode ser considerada, até então, o SOC mais aberto existente. Aberto, diz-se, pois ela aproxima, com velocidade, o processo de classificação por usuários, facilitando a busca e a recuperação de itens informacionais de acordo com experiências e perspectivas de um determinado domínio.

Neste sentido, confere-se a definição dada por Moreiro González (2011, p. 46), ao considerar que a Folksonomia é um "[...] modo de indexação cooperativa que adota palavras-chave para representar a informação de maneira espontânea e livre". O autor ainda completa a definição ao apontar que as folksonomias "[...] beneficiam-se do vocabulário pessoal e da contribuição social para ajudar a encontrar e controlar na web os assuntos mais interessantes" e que "[...] não há uma autoridade central que imponha sua visão, de forma que cada voz tem seu lugar" (MOREIRO GONZÁLEZ, 2011, p. 47-48). Aqui, destaca-se o papel central do usuário, que assume uma posição de classificador de informação e dispensa, assim, o uso de normas e leis, para aplicar ao objeto classificado o seu conhecimento.

Santos e Corrêa (2015, p. 72), ao buscar uma definição de Folksonomia, acreditam que esta é conceituada

> [...] como o resultado do processo de etiquetagem livre (atribuição de etiquetas, palavras-chave) realizada pelos usuários mediante o emprego de termos provenientes de linguagem natural - dispensando o auxílio de vocabulários controlados[20]- em ambientes digitais colaborativos visando indexar recursos informacionais compartilhados em qualquer formato (textos, imagens, áudio, vídeo etc.) para fins de sua representação.

Hjorland (2015, p. 120) define folksonomia como um tipo de taxonomia gerada pelo usuário, com pouca formalidade e precisão se comparadas com as taxonomias tradicionais. Para o autor esse sistema é um exemplo de vocabulário não controlado. Apesar disso, Moreiro González (2011) apontou dois tipos de folksonomias existentes, enquanto uma (folksonomia genérica) mantém o uso de vocabulário não controlado, um outro tipo (folksonomia específica) apresenta termos pré-

[20] "Vocabulário controlado é uma lista restrita e autorizada de palavras ou termos usados para descrever, indexar ou classificar documentos ou informações ... todos os termos em um vocabulário controlado devem ter uma definição não ambígua e não redundante" (HJORLAND, 2015, p. 120, tradução nossa).
"Os vocabulários controlados consistem em termos, palavras da linguagem natural selecionadas como úteis para fins de recuperação pelos designers de vocabulário" (TUDHOPE; NIELSEN, 2006, p. 4, tradução nossa).

determinados ao processo de etiquetagem dos objetos. Observe o esquema a seguir.

Figura 29 - Tipos de folksonomia

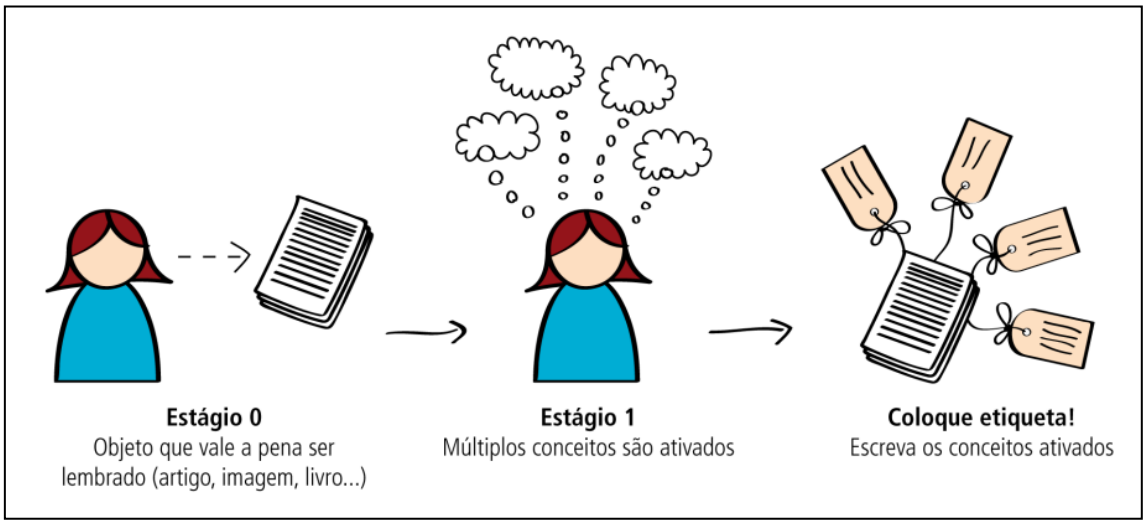
Fonte: O autor (2022) com base em Moreiro González (2011, p. 48).

Ao relacionar a definição de Folksonomia como resultado de um processo de etiquetagem, é imprescindível destacar que este termo deriva do inglês *tagging* e é “uma forma de indexação em que as próprias pessoas, no caso, os usuários da informação, classificam os documentos”, tão logo “[...] o resultado do processo de etiquetagem gera o que vem sendo denominado de folksonomia” (BRANDT, 2009, p. 37-38).

A Folksonomia é um dos tipos de SOC que está situado dentro do ambiente web, mais especificamente dentro da arquitetura da informação social. Esse sistema, livre de regras, facilita o processo classificatório feito por usuários inexperientes, pois o tempo que investem neste processo é inferior ao tempo que eles levariam para utilizar outros sistemas complexos de organização e representação do conhecimento (MOREIRO GONZÁLEZ, 2011).

Não existe um processo estruturado de classificação neste sistema, mas um estágio cognitivo em que o indivíduo descreve o *corpus* de forma a representá-lo para atividades futuras. Tal estágio é interpretado e ilustrado por Brandt (2009) ao se servir de Sinha (2005), identificando três ações necessárias para etiquetagem, ilustradas a seguir:

Figura 30 - Processo cognitivo por trás da etiquetagem



Fonte: (BRANDT, 2009; SINHA, 2005)

No sentido de compreender o processo cognitivo ilustrado na figura, Brandt (2009, p. 65) explica que

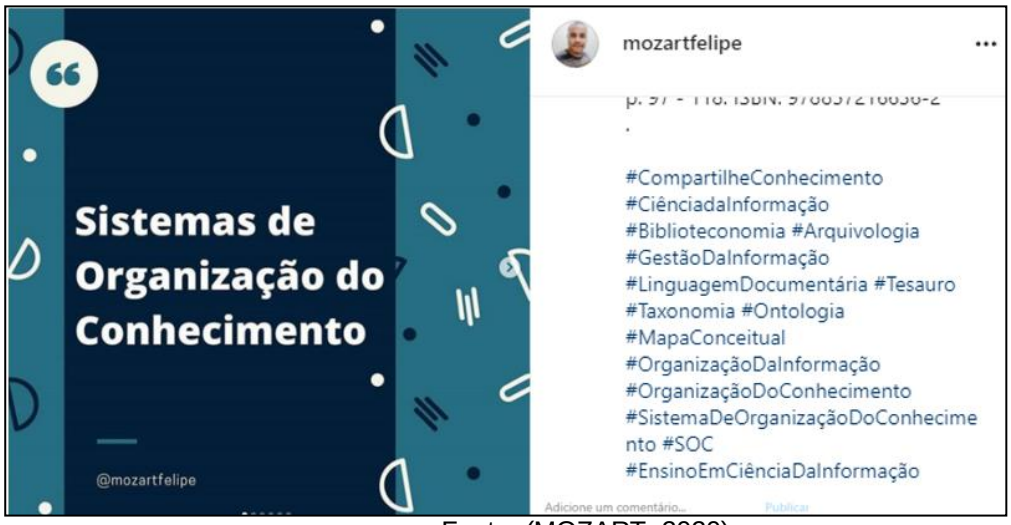
> O estágio 0 corresponde à seleção de um item (objeto de informação) que vale a pena ser recuperado num momento posterior. O próximo estágio é um processo de comparação de similaridade entre o item e os conceitos candidatos a virarem etiquetas. Nessa fase, conceitos semânticos relacionados são ativados: tanto conceitos gerais quanto os mais pessoais a que o objeto remete e ainda características físicas podem ser lembradas, segundos estudos da psicologia cognitiva. A partir do momento que esses conceitos são ativados, externalizá-los, escrevendo-os na forma de etiquetas é bastante simples: não há um processo de filtragem ou julgamento nesse estágio, pode-se usar tantas associações quanto necessário.

Assim, as Folksonomias nos apresentam um conjunto de conceitos já conhecidos, tendo a vantagem de prover da capacidade de associação das palavras às reais experiências dos usuários, não se preocupando, portanto, com uma precisão.

Contudo, esse sistema também representa a desvantagem de causar certa ambiguidade dos termos apresentados pelas etiquetas, pois um mesmo termo pode ser interpretado de formas distintas (MOREIRO GONZÁLEZ, 2011). Isso faz com que a linguagem utilizada pelos usuários, para etiquetar um objeto, seja transformada constantemente, o que faz com que se remeta às reflexões de Dahlberg sobre o uso de linguagens - a autora a diferencia entre linguagens formalizadas e linguagem naturais, e, no caso das Folksonomias, pode-se utilizar os dois tipos de linguagens, embora a linguagem natural seja mais predominante.

A seguir, podem ser observados dois exemplos do uso de etiquetas ou *tags* para representar um conteúdo compartilhado na rede.

Figura 31 - Uso de linguagem formalizada no processo de etiquetagem livre



Fonte: (MOZART, 2020).

Figura 32 - Uso de linguagem natural no processo de etiquetagem livre

Fonte: (BH, 2020).

Percebe-se, então, que, por meio do processo de etiquetagem livre, fica a critério do usuário quais termos são utilizados e qual o objetivo da representação daquele objeto informacional. Na Figura 31, por exemplo, observa-se um conteúdo voltado à comunidade de Ciência da Informação, posto que, evidentemente, o usuário utilizou etiquetas de linguagem formal para representar o objeto, aplicando termos formais da área, como, por exemplo, as etiquetas que representam os cursos de graduação desta Ciência, a saber: #Biblioteconomia; #Arquivologia;

#GestãodaInformação; e ainda termos técnicos desta comunidade, como #LinguagemDocumentária; #OrganizaçãodaInformação; etc.

Já na Figura 32, trata-se de um conteúdo orientado à comunidade específica, pois o usuário utilizou etiquetas de linguagem natural para representar o objeto, empregando as etiquetas que representam as manifestações culturais deste domínio, como: #funk; #bailão; #baile; #bailedagaiola; além de termos mais específicos: #passinhos; #passinhofunk; #passinhodosmaloka.

Percebe-se, portanto, que não há hierarquias para a classificação dos objetos, bem como não são impostas leis ou o controle de singular e plural dos termos utilizados; não existem guias, ou quaisquer limitações para classificar os conteúdos. Assim, a Folksonomia permite que a comunidade de Ciência da Informação, sobretudo no campo de Organização da Informação e do Conhecimento, estude os termos usados por um domínio específico e entenda como eles se comportam, de acordo com as tendências e as comunidades que os empregam (MOREIRO GONZÁLEZ, 2011).

Com a Folksonomia, por meio da indexação colaborativa, é possível organizar e representar um objeto, considerando o domínio ou a comunidade a qual o objeto está direcionado, e o uso de descritores livres eleitos, esse sistema é tido como independente, pois o uso das etiquetas varia entre usuários, grupos de usuários e domínio maiores. Aqui, o usuário é um agente ativo em todo o processo de organização, representação, busca, acesso e uso do conteúdo etiquetado.

Com relação aos demais sistemas a folksonomia é percebida como o mais simples e, também, a que mais possui características que podem contribuir para a construção de outros sistemas de organização do conhecimento, como o desenvolvimento de Tesauros, considerando a rica variedade linguística empregada por distintos domínios; o desenvolvimento de Taxonomias, quanto à classificação por palavras chave em determinados sites; e Ontologias, por permitir o intercâmbio de uma busca e relacioná-la a outros resultados precisos.

Tesauro, Taxonomia e Ontologias são os sistemas alvo de estudo desta tese e serão discutidos a seguir.

3.5 Tesauro

O volume de informações percebido após a Revolução Industrial, século XIX, intensificou a necessidade do desenvolvimento de instrumentos para realizar o controle das informações produzidas na época, tendo, portanto, o tesauro como uma das primeiras iniciativas adotadas como solução.

Em sentido etimológico, o termo “tesauro” deriva do latim *thesaurus*, que se origina do grego *thesaurós*, cuja definição é “estoque de tesouros” ou, simplesmente, “tesouro”. Em um traço histórico do uso do termo, corrobora-se que o termo se originou da Lexicografia, campo dedicado aos estudos dos processos de desenvolvimento de dicionários (CURRÁS, 2010; FELIPE; SILVA NETO, 2019; MACULAN, 2015).

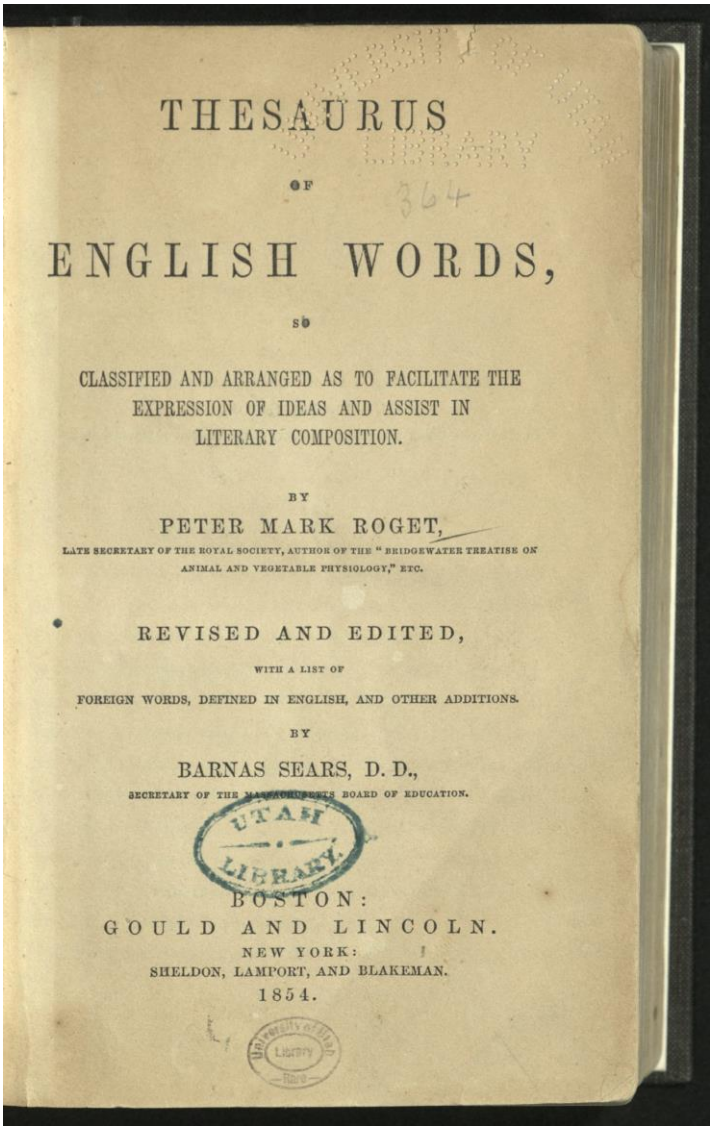
Maculan (2015, p. 129), aportando-se de Gilchrist (1971), ao resgatar os aspectos históricos correspondentes ao uso do termo “tesauro”, aponta que, inicialmente, ele foi utilizado em 1565, por meio da publicação intitulada *Thesaurus Linguae Romanae et Britannicae*, e, em seguida, em 1736, no *Short Oxford Dictionary*. Currás (2010, p. 94), identificou o uso do termo no *Short Oxford Dictionary,* o referenciando como “[...] um armazém de conhecimentos; por exemplo, um dicionário, uma enciclopédia e outras obras semelhantes”.

Historicamente, o uso popular do termo se deu em 1852, com a primeira edição do *Thesaurus of English Words and Phrases*, de Peter Mark Roget. Felipe e Silva Neto (2019, p. 32) apontam que o objetivo de Roget era “[...] exibir as palavras classificadas e arranjadas para facilitar a expressão das ideias e para ajudar na composição literária”. Ainda para os autores:

> Roget considerou a organização de mil seções, ou tópicos, expressos no quadro sinótico de categorias, responsável por especificar as divisões e as subdivisões de cada categoria conceitual existente entre os componentes como um todo. Outra informação que pode ser referenciada é o *Roget’s thesaurus* que deu início ao surgimento do gênero chamado *Thesaurus for English words*. Esse tipo de tesauro expõe a língua inglesa por meio de relações conceituais entre os termos, de forma diferente dos dicionários, mesmo possuindo uma estrutura parecida, no que diz respeito ao suporte. Sua publicação é contínua até o presente momento. (...) O Thesaurus de Roget foi lançado em 29 de abril de 1852, contendo 15.000 palavras. A segunda edição foi publicada em 1853. A primeira edição americana veio em 1854. Durante o último século e meio, várias edições foram lançadas e o Thesaurus de Roget nunca ficou

fora de catálogo. Mais de 32 milhões de cópias foram vendidas em todo o mundo. Seu impacto na escrita é incomensurável (BINGHAM, 2017, *online*, tradução nossa).

Figura 33 - Thesaurus of English Words.



THESAURUS

OF

ENGLISH WORDS,

SO

CLASSIFIED AND ARRANGED AS TO FACILITATE THE EXPRESSION OF IDEAS AND ASSIST IN LITERARY COMPOSITION.

BY

PETER MARK ROGET,

LATE SECRETARY OF THE ROYAL SOCIETY, AUTHOR OF THE "BRIDGEWATER TREATISE ON ANIMAL AND VEGETABLE PHYSIOLOGY," ETC.

REVISED AND EDITED,

WITH A LIST OF

FOREIGN WORDS, DEFINED IN ENGLISH, AND OTHER ADDITIONS.

BY

BARNAS SEARS, D. D.,

SECRETARY OF THE MASSACHUSETTS BOARD OF EDUCATION.

UTAH LIBRARY.

BOSTON:

GOULD AND LINCOLN.

NEW YORK:

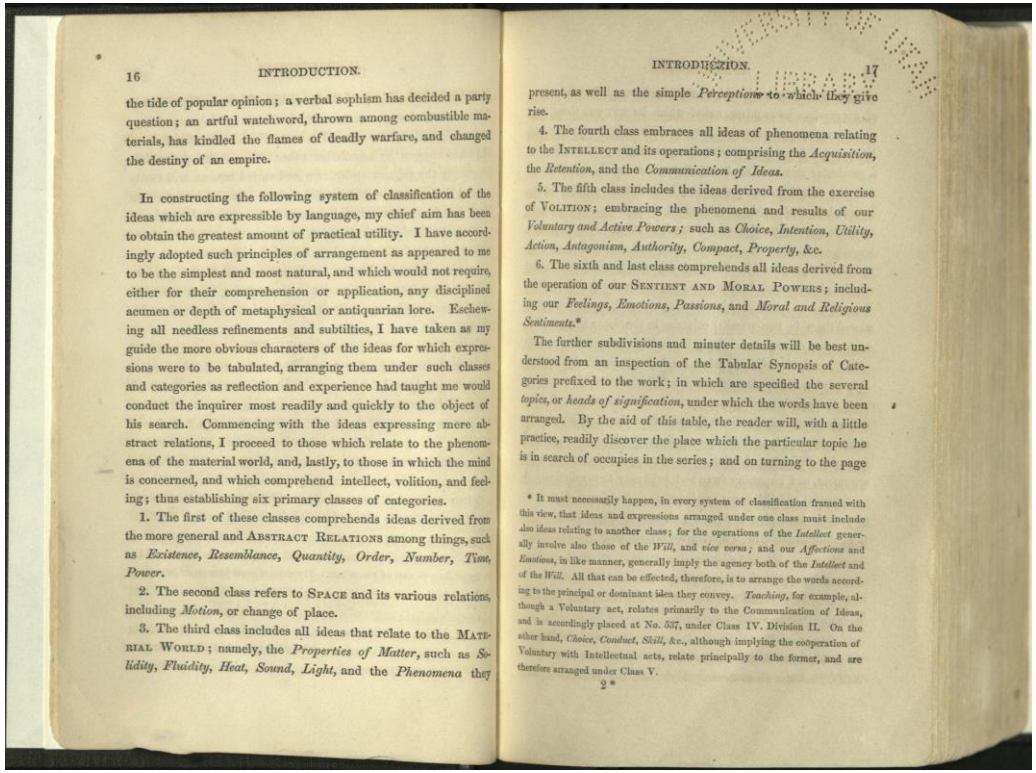
SHELDON, LAMPORT, AND BLAKEMAN.

1854.

University of Utah Library Rare

Fonte: (BINGHAM, 2017, *online*).

Figura 35 - Introduction of Thesaurus of English Words



16 INTRODUCTION.

the tide of popular opinion; a verbal sophism has decided a party question; an artful watchword, thrown among combustible materials, has kindled the flames of deadly warfare, and changed the destiny of an empire.

In constructing the following system of classification of the ideas which are expressible by language, my chief aim has been to obtain the greatest amount of practical utility. I have accordingly adopted such principles of arrangement as appeared to me to be the simplest and most natural, and which would not require, either for their comprehension or application, any disciplined acumen or depth of metaphysical or antiquarian lore. Eschewing all needless refinements and subtilties, I have taken as my guide the more obvious characters of the ideas for which expressions were to be tabulated, arranging them under such classes and categories as reflection and experience had taught me would conduct the inquirer most readily and quickly to the object of his search. Commencing with the ideas expressing mere abstract relations, I proceed to those which relate to the phenomena of the material world, and, lastly, to those in which the mind is concerned, and which comprehend intellect, volition, and feeling; thus establishing six primary classes of categories.

1. The first of these classes comprehends ideas derived from the more general and ABSTRACT RELATIONS among things, such as *Existence, Resemblance, Quantity, Order, Number, Time, Power.*

2. The second class refers to SPACE and its various relations, including *Motion*, or change of place.

3. The third class includes all ideas that relate to the MATERIAL WORLD; namely, the *Properties of Matter*, such as *Solidity, Fluidity, Heat, Sound, Light*, and the *Phenomena* they

INTRODUCTION. 17

present, as well as the simple *Perceptions* to which they give rise.

4. The fourth class embraces all ideas of phenomena relating to the INTELLECT and its operations; comprising the *Acquisition*, the *Retention*, and the *Communication of Ideas.*

5. The fifth class includes the ideas derived from the exercise of VOLITION; embracing the phenomena and results of our *Voluntary and Active Powers*; such as *Choice, Intention, Utility, Action, Antagonism, Authority, Compact, Property*, &c.

6. The sixth and last class comprehends all ideas derived from the operation of our SENTIENT AND MORAL POWERS; including our *Feelings, Emotions, Passions*, and *Moral and Religious Sentiments.**

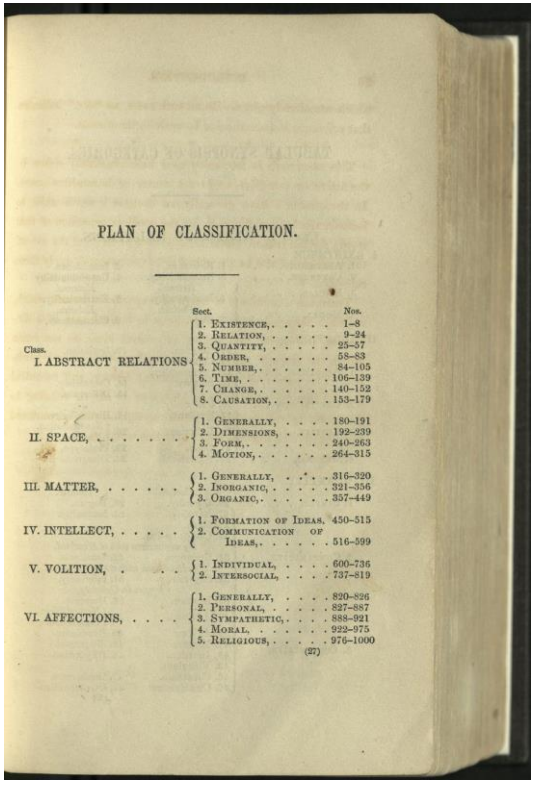
The further subdivisions and minuter details will be best understood from an inspection of the Tabular Synopsis of Categories prefixed to the work; in which are specified the several *topics*, or *heads of signification*, under which the words have been arranged. By the aid of this table, the reader will, with a little practice, readily discover the place which the particular topic he is in search of occupies in the series; and on turning to the page

* It must necessarily happen, in every system of classification framed with this view, that ideas and expressions arranged under one class must include also ideas relating to another class; for the operations of the *Intellect* generally involve also those of the *Will*, and *vice versa*; and our *Affections* and *Emotions*, in like manner, generally imply the agency both of the *Intellect* and of the *Will*. All that can be effected, therefore, is to arrange the words according to the principal or dominant idea they convey. *Teaching*, for example, although a Voluntary act, relates primarily to the Communication of Ideas, and is accordingly placed at No. 537, under Class IV. Division II. On the other hand, *Choice, Conduct, Skill*, &c., although implying the cooperation of Voluntary with Intellectual acts, relate principally to the former, and are therefore arranged under Class V.

2*

Fonte: (BINGHAM, 2017, *online*).

Figura 34 - Plan of Classification of Thesaurus of English Words.



PLAN OF CLASSIFICATION.

Class.		Sect.	Nos.
I. ABSTRACT RELATIONS	1. EXISTENCE,		1–8
	2. RELATION,		9–24
	3. QUANTITY,		25–57
	4. ORDER,		58–83
	5. NUMBER,		84–105
	6. TIME,		106–139
	7. CHANGE,		140–152
	8. CAUSATION,		153–179
II. SPACE,	1. GENERALLY,		180–191
	2. DIMENSIONS,		192–239
	3. FORM,		240–263
	4. MOTION,		264–315
III. MATTER,	1. GENERALLY,		316–320
	2. INORGANIC,		321–356
	3. ORGANIC,		357–449
IV. INTELLECT,	1. FORMATION OF IDEAS,		450–515
	2. COMMUNICATION OF IDEAS,		516–599
V. VOLITION,	1. INDIVIDUAL,		600–736
	2. INTERSOCIAL,		737–819
VI. AFFECTIONS,	1. GENERALLY,		820–826
	2. PERSONAL,		827–887
	3. SYMPATHETIC,		888–921
	4. MORAL,		922–975
	5. RELIGIOUS,		976–1000

(27)

Fonte: (BINGHAM, 2017, *online*).

Tesauro, para a Ciência da Informação, é visto como um sistema de organização orientado à recuperação da informação, e a razão do seu uso foi dada pela substituição das listas de cabeçalhos de assunto. Emília Currás (2010, p. 103), ao tratar da percepção do uso do termo, explana que:

> Os sistemas hierárquicos ou facetados, não mais respondiam de forma adequada às demandas de informação. [...] Era preciso encontrar uma solução. A solução veio daqueles que pensaram que o problema era encontrar, imaginar, novos sistemas de classificação que permitissem maior flexibilidade no tratamento dos temas dos documentos. Observou-se, também, que a ordenação alfabética dos temas apresentava lacunas na hora de intercambiar suas relações. Olhando para o passado, a atenção se ficou naqueles documentos que apresentavam estruturas sistemáticas na ordenação dos conceitos. [...] Dessa forma surgiram os primeiros métodos de classificação utilizando conceitos extraídos dos próprios documentos, sem conexão prévia. O produto resultante recebeu o nome de *thesaurus*...

Ainda para a Ciência da Informação, Moreiro González (2011, p. 63) conceitua o sistema como "[...] lista de descritores (termos controlados) que representam os conceitos de um domínio do conhecimento. E se organiza em estrutura hierárquica com relações semânticas entre si". Conforme a Figura 36, o Tesauro tem os objetivos de:

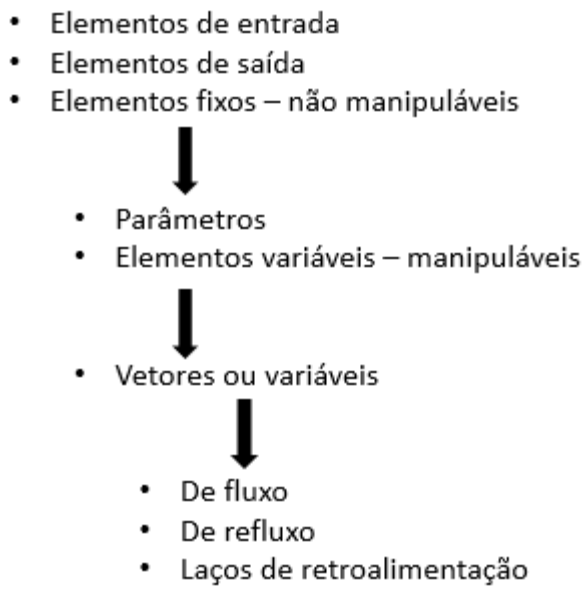
Figura 36 - Objetivos do sistema tesauro.

Fonte: O autor (2022) com base em Moreiro González (2011).

Com o passar dos anos, sendo aprimorado pelo aperfeiçoamento de tecnologias digitais, o sistema de organização tesauro passou por processos de

desenvolvimento distintos, ressignificando, por assim dizer, conceitos e processos. Dito isso, de acordo com a bibliografia consultada, perceber o tesauro nas instâncias tecnológicas requer a identificação de componentes que o constituem como sistema (CURRÁS, 2010, p. 157), são eles:

Figura 37 - Componentes do sistema de tesauro



Fonte: Currás (2010, p. 157)

Em explanação dos componentes identificados acima, Currás (2010) explica que:

- os *elementos de entrada* são responsáveis por alimentar o sistema. São determinados pelas fontes que originam os termos, como: documentos, vocabulários, glossários etc.;

- os *elementos de saída* são vistos como o próprio sistema;

- os *elementos fixos*, que não podem ser manipulados, conformam os parâmetros e são constituídos pelos termos – descritores – que formalizam o sistema;

- os *elementos variáveis*, que podem ser manipulados, são formados por vetores e variáveis determinados pelos vetores de fluxo e de refluxo, e os laços de retroalimentação, responsáveis pela evolução do sistema.

Destaca-se o tesauro como “[...] um sistema altamente complexo, formado por diferentes classes de sistemas e subsistemas” (CURRÁS, 2010, p. 158). Para simplificar, o tesauro integra o sistema principal, os termos e os descritores.

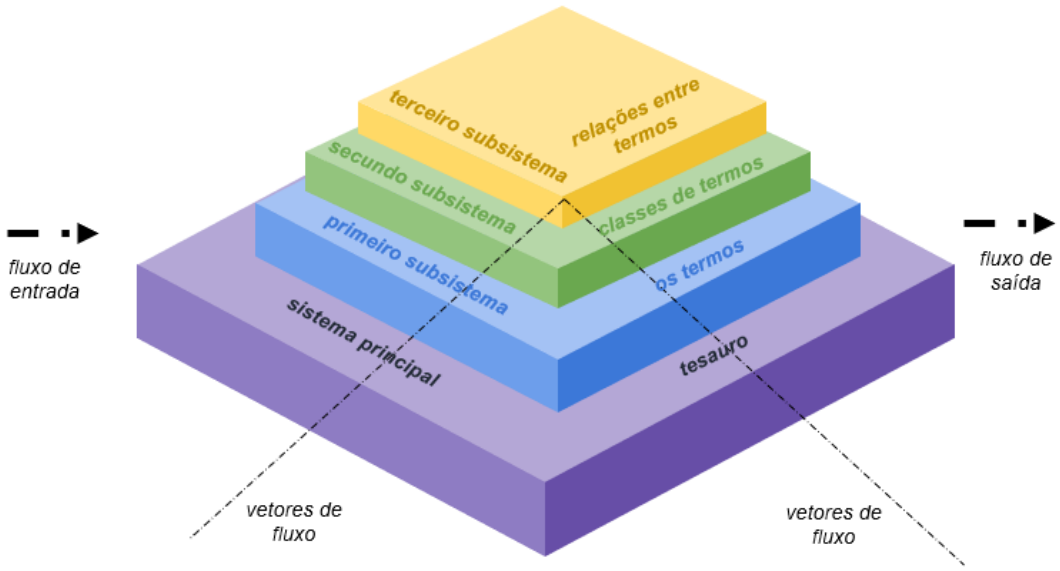
- o sistema principal é o próprio tesauro, que figura os termos, as relações e as influências. É um sistema complexo, resultante dos subsistemas correspondentes. E, pela sua vez, também possui alta complexidade;

- os termos também formarão um sistema parametrizado por classes de termos: principais ou descritores, secundários e sinônimos;

- os descritores se estruturam em famílias formadas por: os termos cabeçalho, genérico, específicos e relacionados.

Em todo o sistema, inclusos os subsistemas, deve-se considerar os elementos variáveis, isto é, os vetores de fluxo, responsáveis por mover o tesauro. Logo, o estudo desse diagrama “[...] permite compreender melhor a forma como se deve proceder na construção do tesauro” (CURRÁS, 2010, p. 160). A Figura 38 reflete a representação do sistema tesauro em conjunto com seus respectivos subsistemas:

Figura 38 - Sistemas e subsistemas de tesauro



Fonte: O autor (2022) baseado em Currás (2010, p. 160)

Em base comum, os tesauros contêm uma série de relacionamentos. Esses relacionamentos são classificados em: de equivalência; hierárquicos; associativos;

partitivos (BRAZ, 2020; CAMPOS; MEDEIROS, 2012; FELIPE, 2016; MOREIRO GONZÁLEZ, 2011; MOREIRO GONZÁLEZ, 2019;).

As *relações de equivalência* estão representadas pelos elementos USE ou UP (usado para) e representam os sinônimos de um determinado termo. As *relações hierárquicas* são as mais evidentes, e constituem-se dos termos genéricos e específicos – utiliza-se os elementos TG (Termo Geral), ou: "aparece com determinado destaque", e TE (Termo Específico). As relações hierárquicas, por sua vez, originam as relações lógicas:

> Para designar o tipo de relação lógica entre os termos, assinala-se a sigla que compreende a posição do termo em evidência e os que se aproximam dele. As siglas que designam relações lógicas são: TG (Termo Genérico, que refere-se ao termo mais abrangente, ou de nível mais alto em relação a outros conceitos, superordenado), TE (Termo Específico, que retrata as subdivisões ou especificações de termos que se referem a conceitos mais amplos, subordinados) e TR (Termo Relacionado, que consistem nas situações de coordenação, ou seja, os conceitos que estão no mesmo nível no sistema hierárquico). (BRAZ, 2020, p. 81)

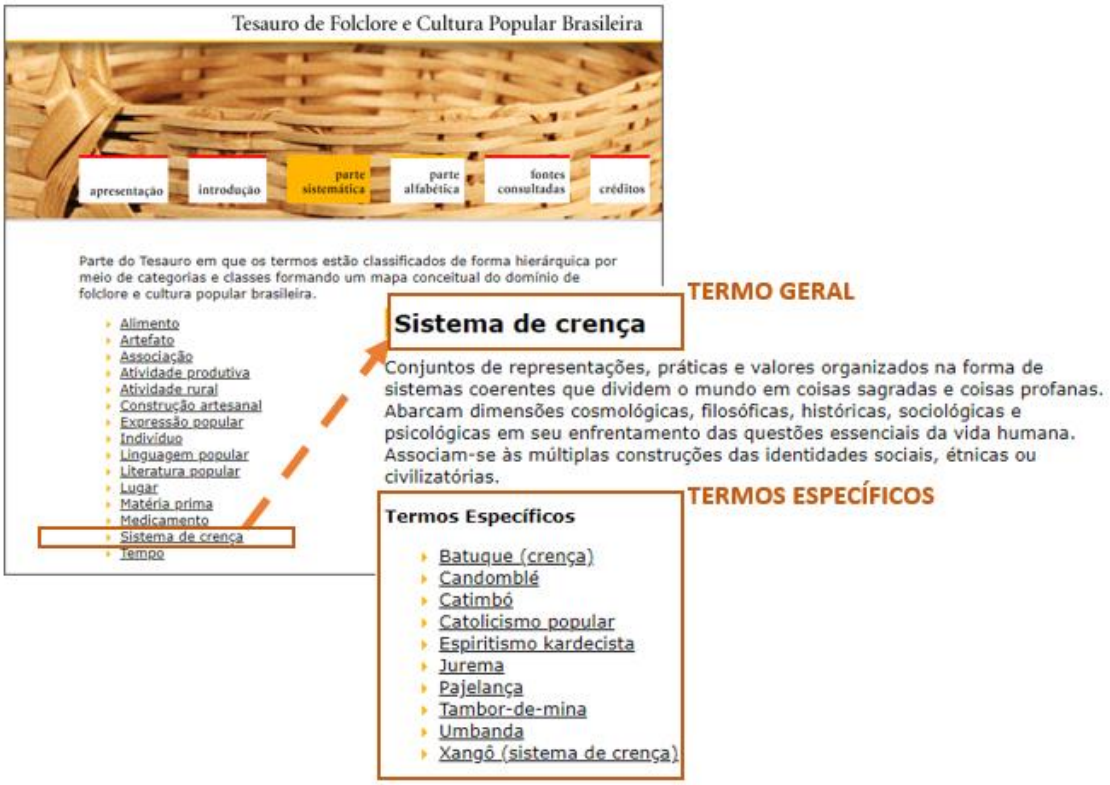
As Figuras a seguir representam a ideia dos tipos de relações; para a relação de equivalência ver Figura 39, para relação hierárquica ver Figura 40.

Figura 39 - Exemplo de relação de equivalência

Fonte: Tesauro de folclore e cultura popular[21]

[21] http://www.cnfcp.gov.br/tesauro/00001969.htm

Figura 40 - Exemplo de relação hierárquica



Fonte: Tesauro de folclore e cultura popular[22]

Os dois exemplos acima ilustraram os tipos de relações de equivalência e de hierarquia extraídos do Tesauro de Folclore e Cultura Popular Brasileira[23]. No exemplo de relação de equivalência, percebe-se o uso dos elementos “usado por” e “use”, relacionando os termos “Ritos de Passagem” – na condição de termo genérico; “Casamento” – na condição de termo específico; e “Núpcias” – na condição de termo equivalente; já no de relação hierárquica, o termo evidenciado, “Sistema de Crença”, é o termo geral, aquele que possui destaque em relação aos termos subsequentes, e, no sentido de facilitar a identificação dos termos, bem

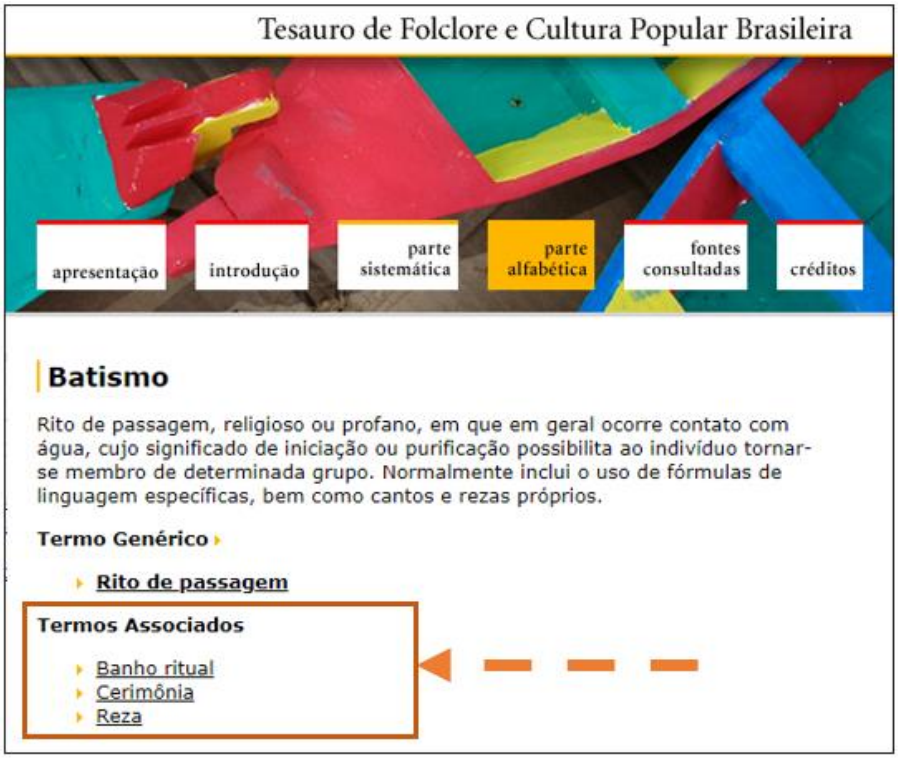
[22] http://www.cnfcp.gov.br/tesauro/00000451.htm

[23] “Este Tesauro é fruto de trabalho sistemático que envolveu uma equipe multidisciplinar, com vasta experiência na área da cultura popular brasileira.
Sua primeira versão, patrocinada pela Unesco, resultou em 2.092 termos, selecionados a partir dos acervos documentais e museológicos do CNFCP. Quando de seu lançamento, em 2004, contava com as categorias: Alimento, Artefato, Associação, Atividade produtiva, Atividade ritual, Indivíduo e Matéria-prima.
A versão atual contempla novas categorias — Literatura oral, Medicamento, Construção artesanal, Sistema de crença, Lugar e Tempo —, além da ampliação das preexistentes Atividade ritual e Atividade produtiva, em que foram inseridas as subcategorias Atividade musical, Atividade narrativa, Prática religiosa e Farmacopéia popular.”
Disponível em: http://www.cnfcp.gov.br/tesauro/apresentacao.html. Acesso em: 24 ago. 2021.

como suas disponibilidades, o tesauro em questão indica os termos específicos que derivam daquele sobreposto. Aqui são termos específicos de “Sistema de Crença”: “batuque”; “candomblé”; “catimbó”; “catolicismo popular”; “xangô”.

Outro tipo de relação encontrada em tesauro é a associativa, que mostra “[...] a possibilidade de categorias estarem, de algum modo, relacionadas, tomando como base o mundo fenomenal” (CAMPOS; MEDEIROS, 2012, p. 102). Exemplifica-se este tipo de relação na figura a seguir.

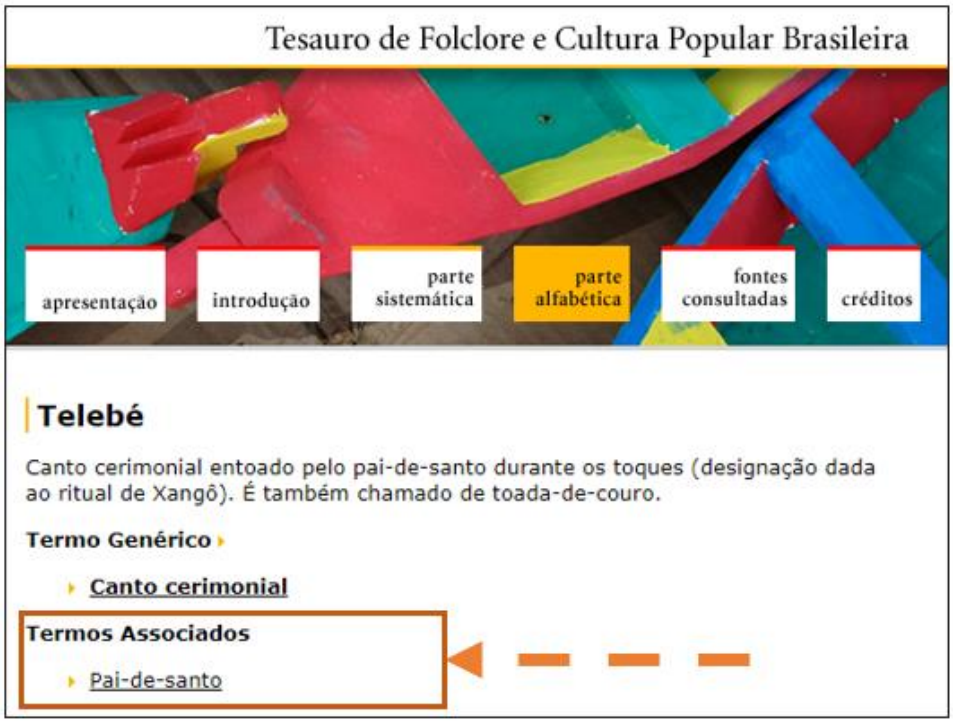
Figura 41 - Exemplo de relação associativa.



Fonte: Tesauro de folclore e cultura popular[24]

[24] http://www.cnfcp.gov.br/tesauro/00001919.htm

Figura 42 - Exemplo (2) de relação associativa



Fonte: Tesauro de folclore e cultura popular[25]

O relacionamento partitivo, ou ontológico, “[...] trata da constituição/formação do objeto” (CAMPOS; MEDEIROS, 2012, p. 102). Esse tipo de relação “[...] ocorre quando há uma significativa conexão semântica entre termos que não estão ligados hierarquicamente” (BRAZ, 2020, p. 83). Como exemplo, tem-se o termo “Seringa”, como objeto de “Vacinação”; e de “Vacinação”, como objeto de “Imunização de doenças”.

Figura 43 - Exemplo de relacionamento ontológico

Fonte: Perobelli (2021)[26]

[25] http://www.cnfcp.gov.br/tesauro/00002172.htm

Na figura 43, percebe-se uma relação entre os três conceitos que a antecederam: seringa, vacinação, imunização de doenças. A relação existente entre esses termos é do tipo ontológica, a qual, como ilustrada na figura, apresenta a ação da imunização à doença COVID-19, por meio de vacinação através da seringa. Não existe uma hierarquia, mas ambas pertencem ao mesmo campo semântico.

A apresentação dos relacionamentos percebidos em uma estrutura, como a de tesauro, permite fundamentar as relações em outros SOC, como o uso do relacionamento hierárquico em taxonomia. A discussão sobre taxonomia é tema da próxima seção.

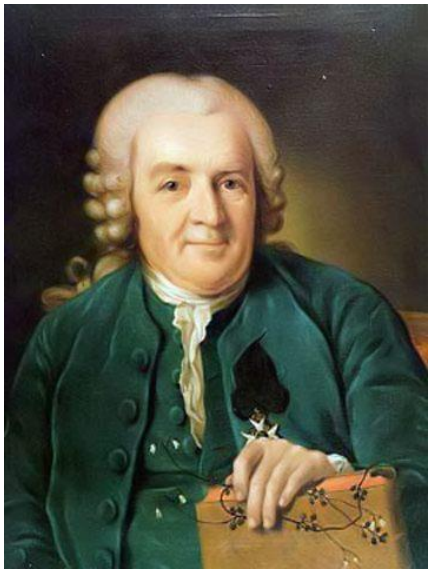
3.6 Taxonomia

Assim como o sistema tesauro derivou de outros campos do conhecimento, como Lexicografia e Terminologia, a taxonomia também teve origem em outro campo do conhecimento: a Biologia. E, posteriormente, integrou-se ao campo da Lógica, desta vez orientada à ordenação de sistemas de recuperação da informação.

O termo foi usado por Carlos Lineu (Figura 44), cientista e médico, em 1735, com a publicação da obra *Systema Naturae* (Figura 45). Foi durante o século XVIII que Lineu classificou e hierarquizou os seres vivos de acordo com suas características. Assim, o cientista os classificou em: reino, filo, classe, ordens, famílias, gêneros e espécies.

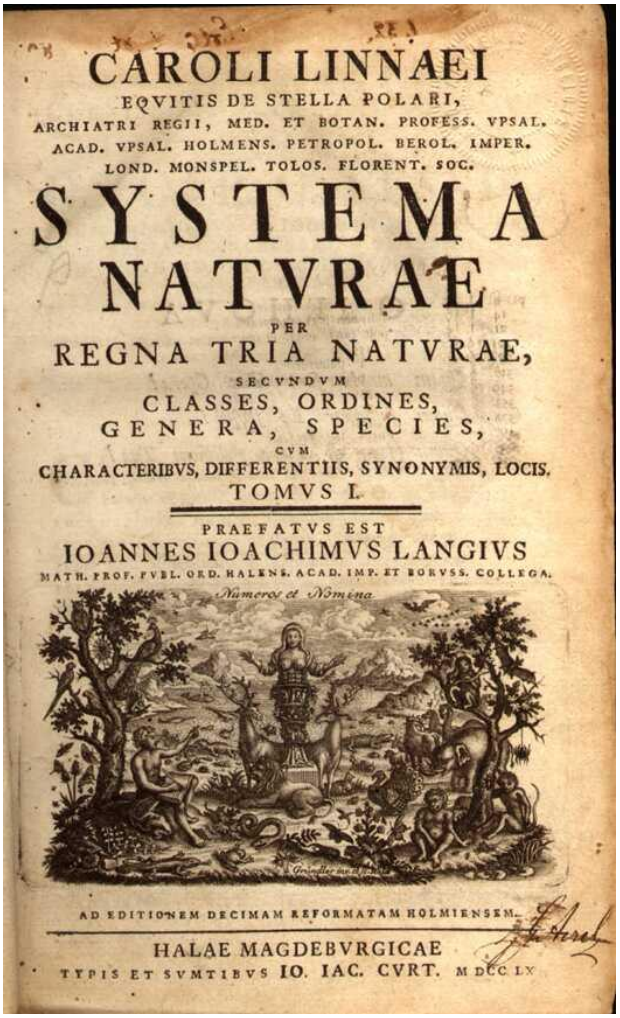
[26] PEROBELLI, A. [Mônica Calazans recebendo vacina]. 1 fotografia, color. São Paulo: G1, 2015. Disponível em: https://g1.globo.com/bemestar/vacina/noticia/2021/01/15/vacinacao-contra-a-covid-19-no-brasil-veja-perguntas-e-respostas.ghtml. Acesso em: 15 abr. 2021.

Figura 44 - Carlos Lineu



Fonte: Biblioteca Ola

Figura 45 - Systema Naturae



Fonte: Wikimedia

Dois séculos após o primeiro uso do termo Taxonomia, por Lineu[27], na Biologia, um outro domínio aportou-se do conceito de taxonomia, o domínio da Educação, mais especificamente a Pedagogia.

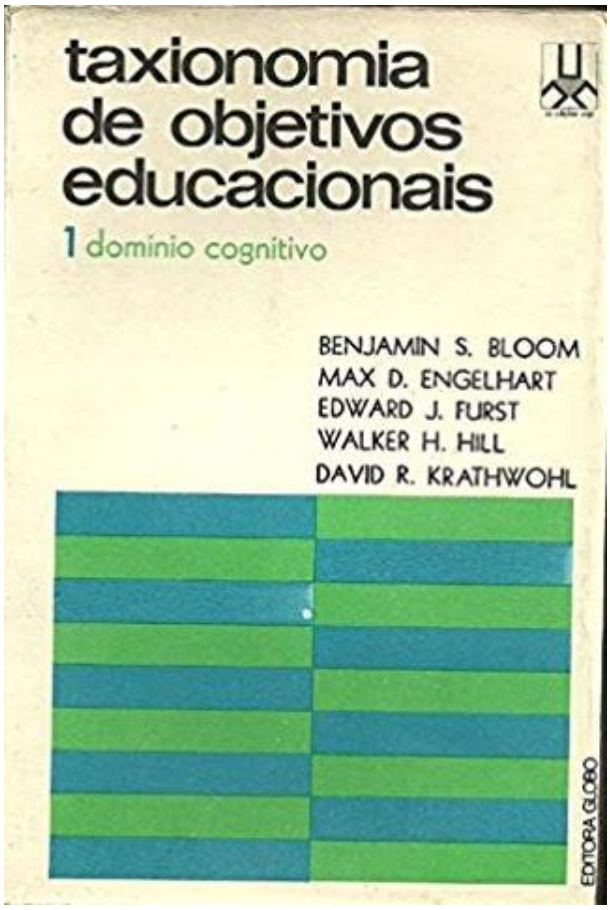
Enquanto na Biologia busca a classificação dos seres, na Pedagogia o termo taxonomia tem seu conceito aplicado na classificação dos objetivos educacionais. O psicólogo e pedagogo Benjamim Bloom[28] (Figura 46), classificou a arte pedagógica em seis ações: avaliação; síntese; análise; aplicação; compreensão e conhecimento.

[27] Imagem 44 disponível em: http://biblioteca-ola.blogspot.com/2016/05/vulgar-de-lineu.html
Imagem 45 disponível em: https://upload.wikimedia.org/wikipedia/commons/a/ac/Systema_Naturae_cover.jpg
[28] Imagem 46 disponível em: https://pt.wikipedia.org/wiki/Benjamin_S._Bloom
Imagem 47 disponível em: http://biblioteca.unicv.edu.cv/upload/2019/11/06/8629.jpg

Figura 47 - Benjamim Bloom

Fonte: Wikipedia

Figura 46 - Taxionomia de objetivos educacionais



Fonte: Biblioteca Unicv

Em Biologia, o termo taxonomia “[...] é a descrição, identificação e classificação dos organismos, individualmente ou em grupo, de alcance global: biotaxonomia, ou especializado: fitotaxonomia e zootaxonomia”, isso porque foi “[...] onde as taxonomias tiveram a função de classificar as espécies em Botânica e Zoologia” (MOREIRO GONZÁLEZ, 2011, p. 51).

Por essa ideia de classificação, este sistema foi adotado nas práticas informáticas, como solução para ordenar e classificar os objetos em sentido lógico. A partir daí, o sistema recebeu maior atenção de profissionais e estudiosos de Ciência da Informação, percebendo este sistema como um meio classificatório, um sistema de classificação (CURRÁS, 2010).

> [...] recentemente, os profissionais da informação e da documentação adotaram o conceito de taxonomia vindo dos informáticos, os quais encontraram nas teorias taxonômicas uma solução para seus problemas classificatórios, principalmente, quando se trata de ordenar unidades conceituais. Estas se extraem dos documentos para melhor estruturar os programas informáticos de indexação e de recuperação da informação. (CURRÁS, 2010, p. 57)

Até meados da primeira década dos anos 2000, era tímida a discussão relacionada ao uso de taxonomias. Contudo, foi nesta mesma época, principalmente devido ao aperfeiçoamento das tecnologias daquele instante, que o sistema taxonômico recebeu maior atenção e aplicabilidade por profissionais da Ciência da Informação – em especial aqueles advindos da organização e representação da informação e do conhecimento.

O enfoque dado à pesquisa e aplicação de taxonomias como sistema classificatório acentuou, consideravelmente, as práticas profissionais e o desenvolvimento de estudos quanto à eficiência desse sistema. Isso deve-se ao fato de que “[...] relacionado, principalmente, aos conceitos de arquitetura da informação, mineração da informação, sempre no contexto da organização do conhecimento em empresas e instituições, levando em conta a web semântica” (CURRÁS, 2010, p. 55).

Observa-se, no entanto, dois conceitos que impulsionaram a aplicação de taxonomias: ordenação e classificação. Aqui, a ordenação está orientada ao

processo físico de dispor um objeto um lugar de pertencimento. A Classificação, neste sistema, é vista como uma atividade complexa, que pode ser vista como aplicação da ordenação orientada por um método (CURRÁS, 2010).

Percebendo destaque naquela definição clássica de taxonomia, e seu aprimoramento a partir dela, Emília Currás (2010, p. 67) abarcou uma série de definições sobre o termo, tais como:

- Taxonomia é campo da ciência que classifica a vida;

- Uma lista organizada de palavras e frases, ou sistema de notação que se usam para iniciar um processo de indexação e recuperação da informação;

- Um esquema de navegação ordenado hierarquicamente;

- Trata de estruturas conceituais adequadas para uso na web semântica;

- Define-se como o que organiza de forma ordenado documentos de temáticas variadas, criando níveis taxonômicos apropriados para o processamento informático;

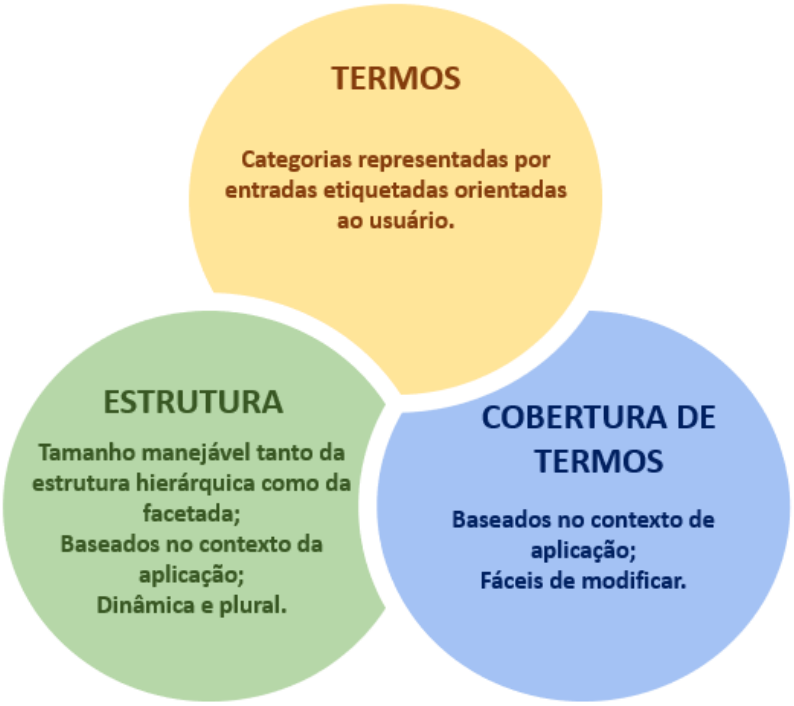
- O que cria ordem no caos de unidades conceituais, advindas de documentos extremamente variados.

Por aferir outros horizontes da taxonomia, em ordenamento conceitual e processual, aprofundar-se-á nas discussões desse sistema, em seu ordenamento conceitual, na seção de “análise de discussão dos resultados”.

De acordo com Moreiro González (2011), uma taxonomia, em sua estrutura básica, é composta por uma organização classificatória hierárquica, com ou sem níveis de profundidades; por rótulos de entradas, sendo estes os termos que compõem a lista, os elementos da estrutura são agrupados tematicamente; e, às vezes, adotam facetas não hierárquicas, as quais dinamizam os enfoques aos assuntos buscados.

A Figura 48 ilustra a estrutura do sistema taxonomia pelos elementos Termos; Estrutura; e Cobertura de Termos.

Figura 48 - Estrutura do sistema taxonomia



Fonte: O autor (2022) baseado em Moreiro González (2011, p. 53).

No sistema taxonômico, diferente dos tesauros – onde os termos precisam de uma justificação documentária, o usuário é o certificador dos termos, ou seja, é necessário que o conteúdo que está sendo apresentado pela taxonomia seja capaz de se adaptar ao conteúdo e às expectativas dos usuários (MOREIRO GONZÁLEZ, 2011). Ainda na estrutura, pode-se estabelecer categorias na classificação, caso se determinem relações de semelhança (princípio da interação) ou de interdependência (princípio da dualidade). No princípio da interação, considera-se um sentido de individualidade, enquanto, que no princípio da dualidade, dar-se-á um sentido de generalidade (CURRÁS, 2010, p. 59).

O aspecto histórico do uso desse sistema reflete nos administradores de empresas, os quais foram os primeiros a buscar auxílio da taxonomia para organização documental. Isso, talvez, porque se deparavam “[...] com uma série de documentos de todo o tipo: contratos, notas fiscais, relatórios técnicos etc., que deveriam ser ordenados de alguma maneira coerente” (CURRÁS, 2010, p. 65), facilitando o processo de busca e recuperação. Além dos administradores de empresas, “[...] foram os informáticos os que recorreram à taxonomia para estudar

sua estrutura e suas características, com a finalidade de aplicá-las à gestão das informações documentárias" (CURRÁS, 2010, p. 65).

Da relação entre a estrutura da taxonomia, os termos e a finalidade de aplicação – considerando a gestão da informação, deve-se considerar quatro aspectos de forte relevância:

Figura 49 - Aspectos relevantes sobre termos empregados na taxonomia.

Fonte: O autor (2022) baseado em Moreiro González (2011, p. 55).

O emprego de termos na construção de SOC é um dos principais critérios para se definir o sistema apropriado para uso, considerando profissionais e usuários. Nesse sentido, pode-se apontar quaisquer sistemas apresentados: classificação, mapas conceituais, folksonomias, tesauros, taxonomias e ontologia, que é discutido a seguir.

3.7 Ontologia

Na década de 1990, presenciou-se o aperfeiçoamento das técnicas de representação de conteúdo, entre elas: tesauros conceituais e verbais; o funcionamento dinâmico do tesauro com mais relacionamentos, buscando melhorar a representação e aumentar sua precisão e eficácia; e as estruturas de navegação de sistemas (MOREIRO GONZALÉZ, 2011) – todos esses apontamentos nos levam a pensar, respectivamente, em tesauros e taxonomias.

Os sistemas de organização do conhecimento, discutidos acima, contribuíram para que outros sistemas surgissem, seja por meio da modelização das estruturas, pelo funcionamento ou pela reunião de atributos que apresentam. Ontologia pode

ser tida como um dos sistemas mais complexos e que reúne características funcionais, estruturais e operacionais dos sistemas apresentados, como tesauro e taxonomia. Não à toa exige-se cada vez mais estudos direcionados à compreensão das diferenças e semelhanças entre SOC.

Assim como tesauro e taxonomia, o sistema ontologia tem origem externa à Ciência da Informação, mas potencializou estudos de organização e representação da informação e do conhecimento, gerando produtos mais sofisticados, para auxiliar nas atividades de busca, acesso e recuperação da informação.

> De uns poucos anos pra cá, vêm surgindo termos novos, dentro dos temas das ciências da informação e da documentação, como por exemplo: domínios (de conhecimento), Web semântica, campos morfológicos..., ontologias. Na realidade trata-se de termos aplicados a conceitos já de muito tempo conhecidos [...]. (CURRÁS, 2010, p. 37)

Para Moreiro González (2011, p. 76), ontologia "[...] é um conceito filosófico que estuda o ser em geral, suas propriedades transcendentais e, mais especificamente, a organização da realidade". O conceito de ontologia ultrapassa as fronteiras dos ramos filosóficos e recebe uma nova percepção no campo da informática, e, mais tarde, na Ciência da Informação. Nascimento (2018) relacionou os três conceitos, reunindo os três campos (filosofia, informática e ciência da informação).

ÁREAS	CONSIDERAÇÕES SOBRE O CONCEITO DE ONTOLOGIA
Ciências Filosóficas	Considerada como o estudo do ser e suas respectivas características e condições de existência.
Ciência da Computação	Considerada como uma especificação formal e explícita de uma conceitualização compartilhada. (GRUBER, 1995).
Ciência da Informação	Considerada como o compartilhamento de conhecimento em uma comunidade por meio de máquinas.
INTEGRAÇÃO DE CONCEITOS DE ONTOLOGIAS ENTRE AS CIÊNCIAS	
É a representação de um conhecimento sobre um determinado domínio, produto da organização da informação realizada por um grupo de indivíduos em computadores, cujo conhecimento é um entendimento consensual para a comunidade o qual é compartilhado.	

Figura 50 - Integração dos conceitos de ontologia.

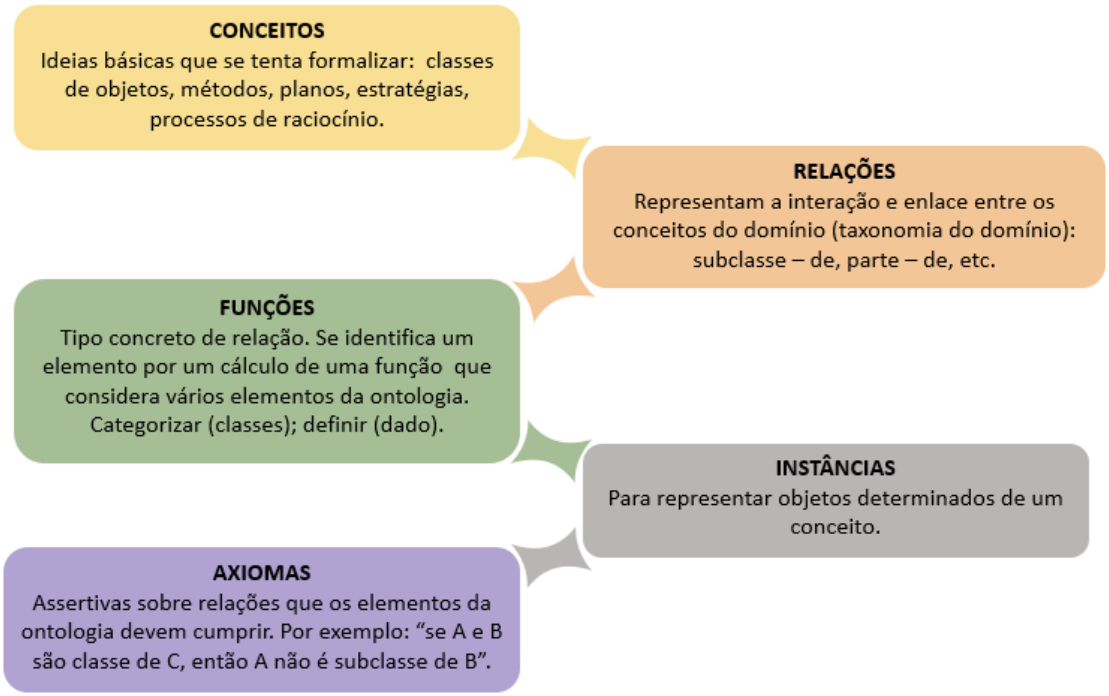
Fonte: Nascimento (2018, p. 44)

O sentido de ontologia adotado nesta tese a percebe como um sistema que "[...] define os termos utilizados para descrever e representar uma área de conhecimento" (MOREIRO GONZÁLEZ, 2011, p. 77), e que, para isso,

> [...] inclui definições de conceitos básicos no domínio e relacionamentos entre eles. [...] As ontologias codificam o conhecimento sobre um domínio, chegando a codificar conhecimento que cobre mais de um domínio. Desta forma, elas fazem com que o conhecimento seja reutilizável. Então, uma ontologia é uma descrição explícita e formal de conceitos em um domínio de discurso (classes, também chamadas conceitos), propriedades de cada conceito, descrevendo várias características e atributos do conceito (*slots* – funções ou propriedades), e restrições sobre os *slots* (facetas – restrições de uma função). Uma ontologia com um conjunto de elementos de classe constitui uma base de conhecimentos. (MOREIRO GONZÁLEZ, 2011, p. 77)

Por uma análise da citação acima, pode-se apontar, portanto, que os componentes básicos para a construção de uma ontologia são conceitos e relações. Além desses três componentes, ontologias contam com funções, instâncias e axiomas. A definição de cada componente citado está apresentada a seguir.

Figura 51 - Componentes de uma ontologia.



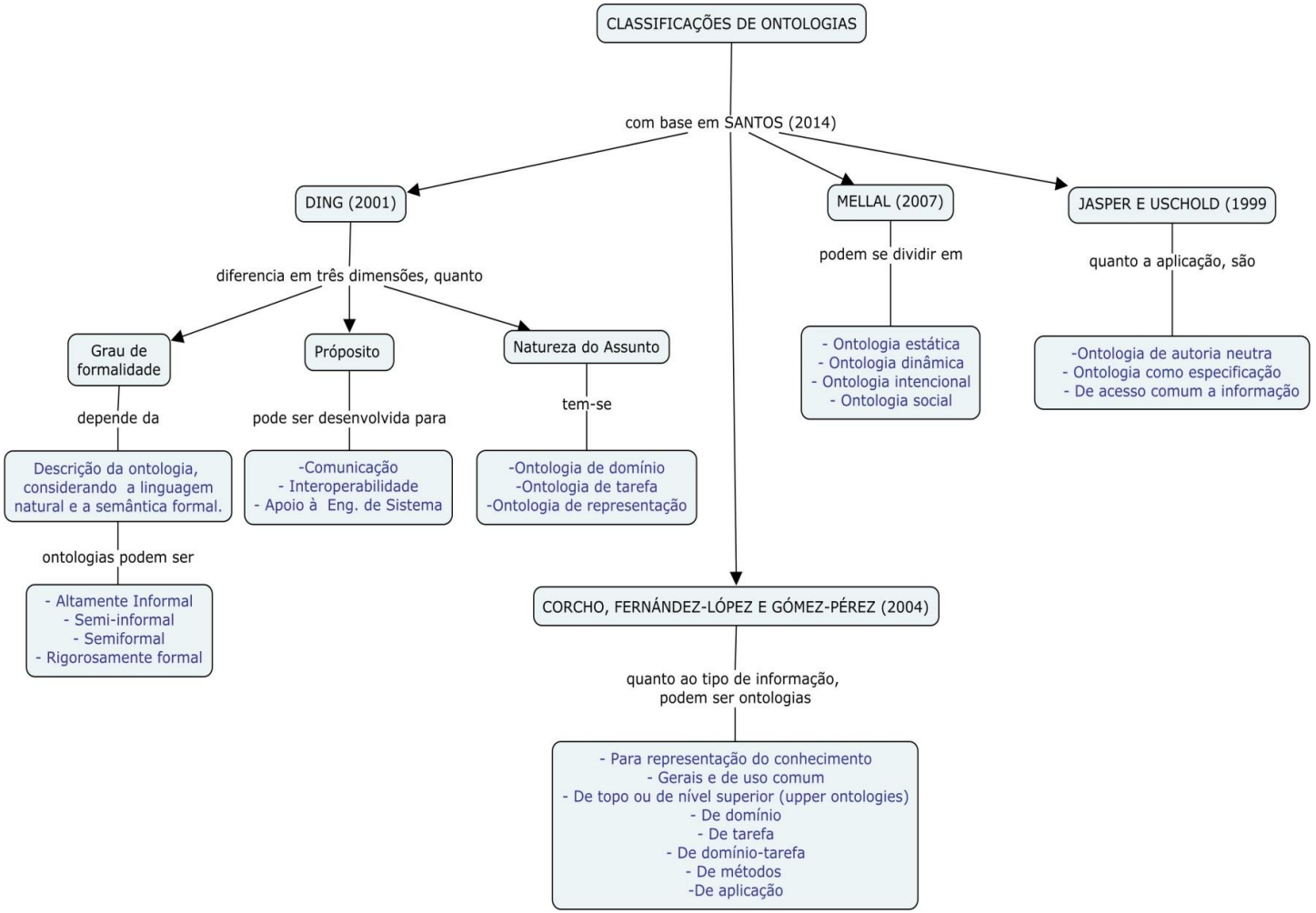
Fonte: O autor (2022) com base em (GRUBER, 1995) citado por Moreiro González (2011, p.78).

De acordo com o grau de complexidade com que se detalham os componentes acima, e com que se exploram os relacionamentos e suas interfaces –

os quais contribuem para determinação de axiomas precisos, pode-se originar ontologias dos mais diversos tipos. Podendo ser simples, comparando-se a taxonomias de domínio; ou mais elaboradas, com relações e interoperações entre conceitos mais robustas; ou, ainda, de maior complexidade.

Nascimento (2018) aportou-se de concepções bibliográficas quanto à classificação de ontologias e elaborou um mapa identificando a diferença na construção desse sistema – grau de formalidade, propósito e natureza (DING, 2001); quanto ao tipo de informação – ontologias gerais, de domínio, de tarefa etc (CORCHO; FERNÁNDEZ-LÓPEZ; GÓMEZ-PÉREZ, 2003); quanto a estrutura – ontologia estática, dinâmica, intencional e social (MELLAL, 2007); e quanto à aplicação, sendo elas: de autoria neutra, de especificação, de acesso (JASPER; USCHOLD, 1999). O mapa da classificação desse sistema está apresentado a seguir.

Figura 52 - Classificações de Ontologia



Fonte: Nascimento (2018, p. 55) com base em Santos (2014).

Nascimento (2018) ressalta que novas buscas sobre esse sistema podem refletir outras classificações, estejam elas baseadas nos níveis de formalidade, em suas aplicações ou funções, em outros tipos de contextos, e, até mesmo, em outras esferas nas quais a utilizem. Para isso, “[...] as ontologias contam com diferentes mecanismos e fases de construção, assim como muitas aplicações e publicações que asseguram a validade de todas as metodologias propostas” (MOREIRO GONZÁLEZ, 2011, p. 85).

Em linhas gerais, é percebido, por meio da publicação de estudos que objetivam desenvolver modelos ou métodos para desenvolvimento de SOC, que a adoção de processos para determinar a estrutura dos sistemas necessita de uma análise e de determinação do foco, dos objetivos, dos usuários do sistema, da necessidade técnica de funcionamento, dos requisitos para operação, dos suportes para armazenamento, recuperação e acesso à informação e ao conhecimento apresentados por esses sistemas. Considera-se, ainda, os aprimoramentos tecnológicos nos aperfeiçoamentos de SOC. Organização, representação, gestão e tecnologia são, portanto, aliadas para desenvolvimento de SOC.

Logo, para além dos processos de desenvolvimento de SOC, os conceitos desses sistemas também foram aperfeiçoados, em razão das tecnologias digitais. Assim, na seção a seguir, estão mapeados os conceitos e processos dos sistemas: tesauro, taxonomia e ontologia, em razão dos estudos elaborados nos últimos anos e publicados nas fontes bibliográficas consultadas, conforme visto no Quadro 2 (páginas 27 e 28).

4 ANÁLISE E DISCUSSÃO DOS RESULTADOS

O termo “Sistemas de Organização do Conhecimento” é considerado genérico, sendo usado para se referir a uma amplitude de sistemas (MAZZOCCHI, 2018), tais como aqueles apresentados na seção 3.

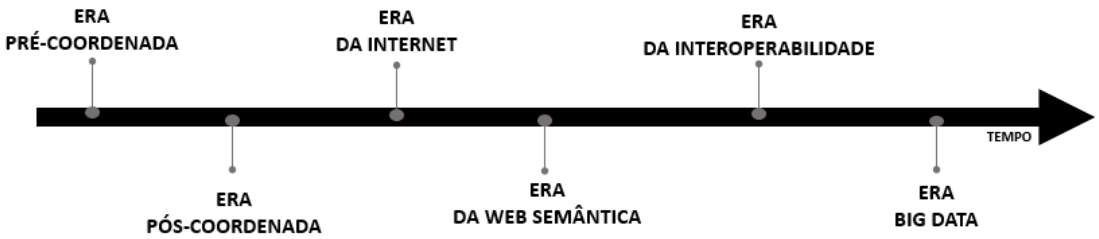
Esta seção busca, portanto, refletir sobre as contribuições das tecnologias digitais no aperfeiçoamento dos SOC: tesauro, taxonomia e ontologia.

4.1 Itinerário evolutivo do desenvolvimento de SOC: interoperabilidade e big data

Antes de partir para a reflexão sobre tais contribuições, é imprescindível caracterizar SOC de acordo com seu itinerário evolutivo, dadas as propriedades particulares dos sistemas, considerando que eles são vistos como esquemas para organização, representação, gestão e recuperação de informações. Esse itinerário pode ser visto em Vickery (2010), ao abordar os estágios de desenvolvimento dos SOC: era pré-coordenada; pós-coordenada; da internet; e da web semântica.

Contudo, ao visualizar o cenário social tecnológico atual, é perceptível que as propriedades classificatórias das eras mencionadas por Vickery (2010) ampliaram-se e alçaram novos horizontes. Logo, é necessário apontar novas eras para somar a este itinerário evolutivo. Esta tese atualiza, também, a inserção de dois estágios no desenvolvimento de SOC. Observe a figura:

Figura 53 - Itinerário evolutivo do desenvolvimento de SOC



Fonte: O autor (2022) com base em Vickery (2010).

Essas eras não se especificam por um intervalo temporal-espacial, mas acompanham as propriedades digitais que surgem e se aprimoram em prol das

melhorias de aplicação, desenvolvimento e uso de sistemas de organização do conhecimento. De acordo com Vickery (2010, *online*):

- A **Era Pré-coordenada** diz respeito à fase em que os SOC assumiram estruturas estáticas, onde as entradas de assuntos podem ser feitas por termos individuais ou termos combinados, como catálogo de fichas;

- A **Era Pós-coordenada** é orientada por SOC que admitiram a forma de elementos capazes de manipulação. Neste estágio, os SOC são dinâmicos.

- A **Era da Internet** aborda três tipos de pesquisas *online* como: a classificação hierárquica e facetada, a referência de citação de um item para outro e o índice do mecanismo de busca. Os SOC são dinâmicos e orientados às estruturas navegacionais.

- A **Era Web Semântica** traz a concepção de criar SOC que possam ser usados por agentes de *softwares*. Potencializa pesquisas e resultados por meio do uso de relações semânticas. Conforme o autor:

> Para que essa pesquisa seja 'significativa', ela deve fazer uso de relações semânticas e, além disso, essas relações devem ser explícitas no KOS. O software de computador só pode 'saber' que uma relação semântica particular existe entre dois termos se tiver sido explicitamente dito assim, ou pelo menos fornecido com dados e procedimentos de inferência pelos quais ele pode 'deduzir' o fato. (VICKERY, 2010, *online*, tradução nossa)

Caracterizadas as quatro eras apontadas por Vickery (2010), as duas outras, incorporadas ao itinerário do desenvolvimento dos SOC, são: **Era da Interoperabilidade** e **Era Big Data**, as quais estão descritas, sequencialmente, abaixo.

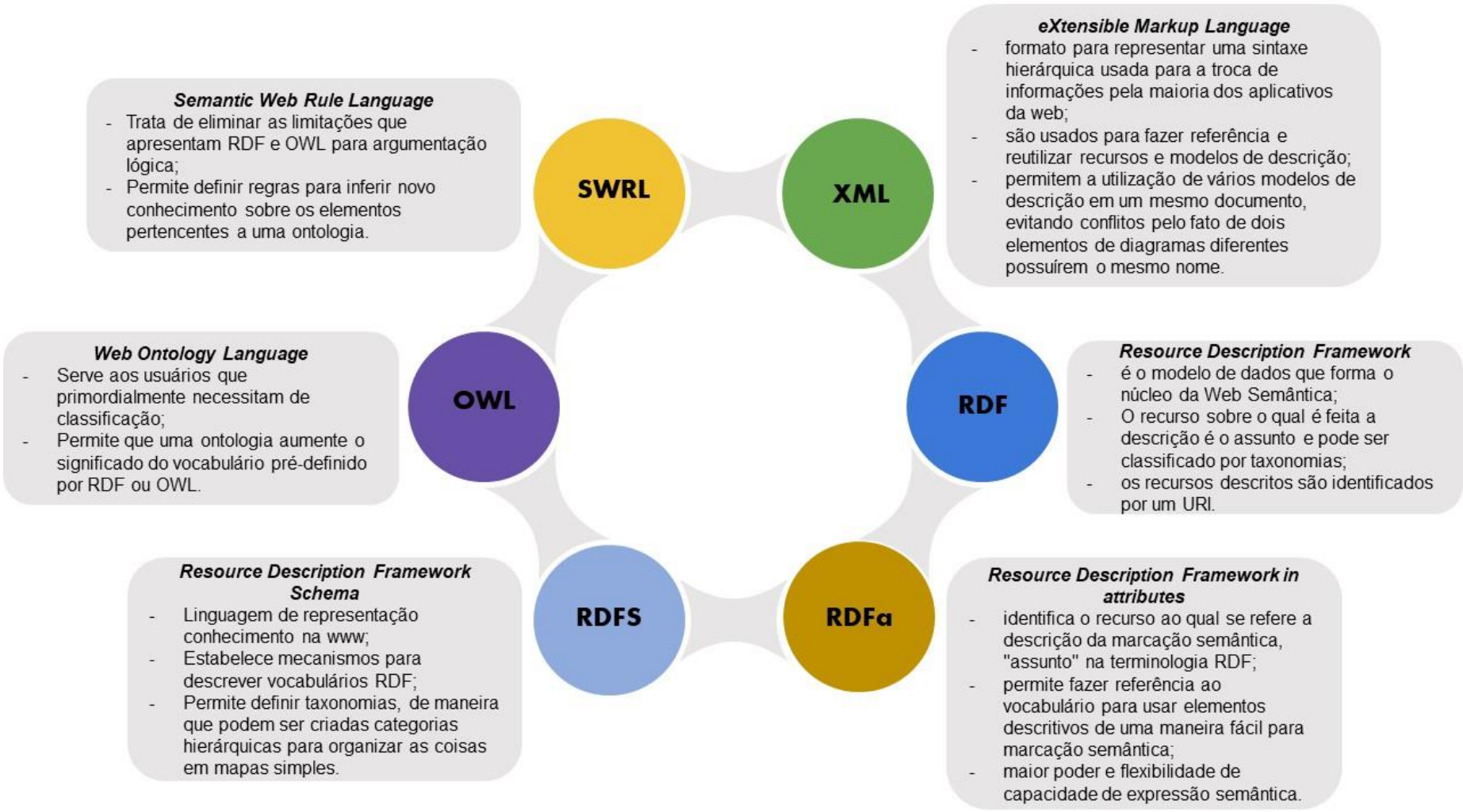
- A **Era da Interoperabilidade**, posterior à da Web Semântica, herdou os desafios daquele estágio, causando uma reestruturação nos modelos de desenvolvimento de SOC, principalmente nos diferentes padrões de representação.

Conforme descrito por Vickery (2010), a era da web semântica propôs que os SOC pudessem ser usados por agentes de *softwares* por meio do uso de relações semânticas. Logo, para que isso ocorra, a web semântica necessita que os dados possam ser legíveis por agentes de *softwares*, os quais também requerem a necessidade de padrões que permitam a representação de informações em modelos interoperáveis, como: *XML* – linguagem de marcação; *RDF* – para representação de metadados; *RDF Schema* – representação do vocabulário em gráficos RDF e

relacioná-lo utilizando estruturas classificatórias, como taxonomias; e OWL – vocabulário padrão para representação de ontologias (DIÉZ *et al.*, 2010). A reunião dessas tecnologias, RDF, RDF Schema e OWL, se complementam e resultam em "[...] uma semântica bem definida para uso na execução de inferências por meio de regras lógicas de raciocínio" (PASTOR SÁNCHEZ, 2013, p. 62, tradução nossa).

A Figura 54 a seguir apresenta a reunião dessas tecnologias para representação da informação em modelos interoperáveis.

Figura 54 - Tecnologias para representação da informação em modelos interoperáveis



Fonte: O autor (2022) com base em Moreiro González (2011) e Pastor Sánchez (2013).

Existem pesquisas que buscam discutir o processo de interoperabilidade entre SOC, como a proposta de Gayo Diallo (2014), no artigo intitulado: *An effective method of large scale ontology matching*, em que apresentam a projeção e a implementação de um método para interoperar ontologias. Ressalta-se que o sistema proposto pelos autores participou da campanha *Ontology Aligment Evaluation Initiative*, nos anos 2012 e 2013, sendo considerado entre os melhores para interoperabilização de ontologia (DIALLO, 2014, p. 1).

- A **Era do Big Data**[29] é vista como um fenômeno que potencializou o uso de técnicas sofisticadas de exploração de dados, tendo isso como uma aceleração das descobertas em determinados campos do conhecimento. Podem-se citar, como exemplos dessas técnicas, o aprendizado de máquinas e a mineração de texto. Esses exemplos vêm junto aos avanços recentes nas técnicas de processamento de dados, que originou algoritmos de aprendizado de máquinas capazes de processar grandes volumes de dados (IBEKWE-SANJUAN; BOWKER, 2017).

> A lista de aplicativos em que algoritmos de big data estão sendo usados agora está em constante crescimento, desde a busca e recuperação mais convencionais até sistemas de recomendação, como os da Amazon e Netflix, que sugerem o que as pessoas podem querer ler a seguir ou assistir a seguir com base nos cliques e compras; à publicidade direcionada; à corretagem em bolsa de valores; para detecção de tendências, mineração de opinião e visualização de informações. Embora não sejam infalíveis, esses algoritmos atingiram um nível de desempenho aceitável para os humanos. Além disso, eles são programados para trabalhar em segundo plano de uma maneira não intrusiva, coletando dados silenciosamente e os processando para fornecer aos usuários sugestões e recomendações que podem rivalizar com as de um bibliotecário humano ou especialista em Organização do Conhecimento. (IBEKWE-SANJUAN; BOWKER, 2017, p. 188, tradução nossa)

Os autores supracitados argumentam, ainda, que, ao considerar Organização do Conhecimento como um campo, “[...] é preciso se adaptar à natureza mutável da produção nas ciências sociais e naturais, na medida em que estas, por sua vez, estão sendo afetadas pelo advento do Big Data”. Constatam, igualmente, que “[...] o campo de Organização do Conhecimento” precisa passar de estabelecer o apodítico (aquilo que conhecemos para sempre) para se adaptar ao novo mundo do

[29] “Há um consenso, pelo menos do ponto de vista físico, que Big Data representa volumes de dados de tal forma que os algoritmos de banco de dados tradicionais são incapazes de lidar com isso e que requer infraestruturas de computador mais robustas e distribuídas e algoritmos” (IBEKWE-SANJUAN; BOWKER, 2017, p. 192, tradução nossa).

conhecimento científico social e natural, criando esquemas maximamente flexíveis. (IBEKWE-SANJUAN; BOWKER, 2017, p. 195, tradução nossa).

A implicação de Big Data, relacionada aos estágios de desenvolvimento dos SOC, traz consigo desafios para a Ciência da Informação e, entre eles, em específico, para Organização do Conhecimento o dever de "[...] reinventar-se em um ecossistema de informações cheio de algoritmos que processam dados continuamente e fornecem conteúdo digital adaptado aos perfis dos usuários" (IBEKWE-SANJUAN; BOWKER, 2017, p. 196, tradução nossa). Esse desafio vai de encontro ao extremo da linha evolutiva; se, em um extremo, encontra-se a era Big Data, no outro extremo tem-se a era Pré-coordenada. Ou seja, enquanto esta é caracterizada por sistemas estáticos, regidos por padrões, sistemas engessados e inflexíveis, aquela é caracterizada por sistemas altamente flexíveis, dispensando padronizações rígidas e tornando os sistemas mais dinâmicos, deixando-os mais próximos à comunidade de usuários.

Para que isso aconteça, é necessária a relação entre estudiosos e profissionais de Organização do Conhecimento, e estudiosos e profissionais de outras áreas do conhecimento, como Tecnologia da Informação e Comunicação. Uma outra visão relevante é refletir sobre os currículos de Biblioteconomia, Arquivologia, Documentação, Gestão e Ciência da Informação, para integrar disciplinas de eixos tecnológicos, como análise de dados, por exemplo. (IBEKWE-SANJUAN; BOWKER, 2017).

As concepções acima, apesar de serem retratadas em literatura internacional, valem também para reflexões nacionais. Apontam-se, aqui, as considerações de Coneglian, Gonçalvez e Satarém Segundo (2017, p. 127), ao enfatizarem que

> [...] o Big Data caracteriza-se essencialmente pelo volume dos dados gerados e disponíveis, pela velocidade com que estes devem ser tratados e apresentados e pela variedade de fontes onde eles se encontram. Essa conjuntura tem provocado uma revolução no modo como as análises de dados estão ocorrendo, tanto no âmbito organizacional, quanto no acadêmico.

Ao retornar para a reflexão da formação e atuação do profissional da informação, considerando a inserção de disciplinas de eixos tecnológicos, utilizando “Análise de Dados” como exemplo, pode-se apontar os seguintes papéis do profissional da informação no processo de análise de dados, um dos processos implicativos do Big Data:

Figura 55 - Processos da análise de dados e o papel do profissional da informação

Fonte: O autor (2022) com base em Coneglian, Gonçalvez e Satarém Segundo (2017, p. 137).

Os processos de organização e representação da informação e do conhecimento na base de análise de dados, compreendendo suas respectivas atividades, são fundamentais para o profissional da informação executar cada fase. Os apontamentos evidenciam-se em Coneglian, Gonçalvez e Satarém Segundo (2017, p. 138).

- *Fontes*: os profissionais devem estar aptos a conduzir a seleção de fontes que terão validades no ato da organização e representação final para os gestores;

- *Captura e armazenamento*: deve ser feita pelos profissionais envolvidos e ser concluída de forma que os dados estejam organizados para a próxima fase;

- *Processamento e fusão*: Aqui, o profissional atua na estruturação dos dados, para possibilitar a fusão da informação. Essa atuação deve-se ao fato de possuir “[...] capacitação para identificar estruturas capazes de converter uma massa de

dados em informações estruturadas como, por exemplo, estruturas de chave-valor e metadados" (CONEGLIAN; GONÇALVEZ; SATARÉM SEGUNDO, 2017, p. 138);

- *Acesso*: essa fase permeia discussões com os gestores, objetivando que os profissionais da informação contribuam na definição dos dados a serem tratados;

- *Análise*: aqui, o profissional da informação é um ator mediador e participa para direcionar os rumos e as variáveis consideradas na análise, isso porque "[...] o profissional da computação é evidentemente o responsável pela construção de algoritmos e pela definição das estatísticas utilizadas" (CONEGLIAN; GONÇALVEZ; SATARÉM SEGUNDO, 2017, p. 138);

- *Exposição*: fase de apresentar de forma bem-organizada e representada os resultados obtidos. É dele a principal função nesta última fase, pois ao assumir a mediação entre profissionais de computação e o gestor, é capaz de apresentar com propriedade os resultados alcançados por meio da análise de dados;

Os exemplos citados introdutoriamente nesta era, como aprendizado de máquinas e mineração de textos, aperfeiçoaram, diretamente, os modelos de construção de SOC como tesauro, taxonomia e ontologia, por exemplo. Esses sistemas serão discutidos a seguir no contexto dos objetivos em que configuram esta tese.

4.2 Tesauros: conceitos e processos

Após consulta e extração das produções publicadas nas fontes bibliográficas – Quadro 2, foram mapeadas as concepções conceituais e processuais apresentadas pelos autores devidamente referenciados. Essas concepções foram organizadas em três quadros, correspondendo, cada um, a um SOC específico. O quadro 7, apresentado a seguir, é sobre tesauro.

Quadro 7 - Concepções conceituais sobre tesauro.

ANO	AUTOR(A)(ES)(AS)	CONCEPÇÕES CONCEITUAIS
2010	MEDEIROS; CAMPOS	age como controlador terminológico para tradução de uma linguagem natural numa linguagem restrita; é um vocabulário controlado e dinâmico de termos que se relacionam semântica e genericamente, cobrindo um domínio

		específico do conhecimento.
2012	OLIVERIA; ARAÚJO.	são instrumentos imprescindíveis à sistematização e disseminação do conhecimento humano.
2012	KROEFF; LEONETTI	é uma lista estruturada de termos associados, empregada por analistas de informação e indexadores, para descrever um documento com a desejada especificidade, em nível de entradas, e para permitir aos pesquisadores recuperação da informação que procura.
2012	CÓRDOVA; LAZARO	é uma linguagem pós-coordenada com estrutura combinatória, vocabulário controlado e especializado por natureza; embora, segundo muitos autores, não tenha nascido com a ideia de pós-coordenação. Ao contrário, surge com a noção básica de cabeçalhos de assuntos, que é uma linguagem pré-coordenada [...] se consolida com o desenvolvimento tecnológico e, principalmente, da informática, com o surgimento dos computadores; o tesauro é um tipo de vocabulário controlado que facilita a comunicação e a transferência de informações, ao eliminar a ambiguidade da linguagem natural e livre em que um documento é escrito; portanto, enfatiza conceitos sobre palavras ... então podemos concluir que o tesauro é uma espécie de linguagem documental que interage com as ideias do autor (documento), do analista da informação documental (indexador) e das necessidades de informação das pessoas (usuários).[30]

[30] es un lenguaje postcoordinado de estructura combinatoria, vocabulario controlado y especializado por naturaleza; pese a que, según muchos autores, no nació con la idea de postcoordinación. Surge más bien con la noción básica de los encabezamientos de materia, que es un lenguaje precoordinado...se consolida con el desarrollo tecnológico y, particularmente, de la informática, con la aparición de las computadoras; el tesauro es un tipo de vocabulario controlado que facilita la comunicación y la transferencia de información, al eliminar la ambigüedad del lenguaje natural y libre en que está escrito un documento; por lo tanto, enfatiza los conceptos sobre las palabras... Entonces podemos concluir que el tesauro es un tipo de lenguaje documental que hace interactuar las ideas del autor (documento), el analista de información documental (indizador) y las necesidades de información de las personas (usuarios).

2012	MARTÍNEZ-GONZÁLEZ; MUÑOZ NIETO	Thesauri são ferramentas conceituais (KOS, Knowledge Organization Systems) que permitem modelar o vocabulário de um domínio; em formato eletrônico exige a disponibilização de ferramentas que permitem a construção e edição dos tesauros, ou a consulta e utilização a partir de várias aplicações. Nesse contexto, há a necessidade, comum no processamento de computadores, de pacotes de software genéricos que facilitem seu trabalho para desenvolvedores de aplicativos.[31]
2013	GONZALEZ AGUILAR	O tesauro tem como objetivo fornecer um vocabulário relacional controlado usado principalmente para indexar e recuperar informações armazenadas em bancos de dados; Um dicionário de sinônimos é precisamente definido como uma lista de autoridade fornecida por uma estrutura geral de rede semântica; A importação e exportação podem ser feitas em XML, RDF ou SQL o que garante a interoperabilidade e integridade estrutural do tesauro.[32]

[31] Los tesauros son herramientas conceptuales (KOS, Knowledge Organization Systems) que permiten modelar el vocabulario de un domínio; en formato electrónico requiere la disponibilidad de herramientas que permitan construir y editar tesauros, o consultarlos y utilizarlos desde aplicaciones diversas. En este contexto se sitúa la necesidad, habitual en los tratamientos informáticos, de paquetes de software genéricos que faciliten su labor a los desarrolladores de aplicaciones.

[32] El tesauro tiene como objetivo proporcionar un vocabulario controlado relacional utilizado principalmente para indexar y recuperar la información almacenada en bases de datos; Se define con precisión un tesauro como una lista de autoridad próvida de uma estructura de red semántica general; La importación y la exportación se puede hacer en XML, RDF o SQL lo que garantiza la interoperabilidad y la integridad estructural del tesauro.

2013	NAUMIS PEÑA; MORÁN REYES; MORALES MONROY	O tesauro evoluiu como uma ferramenta de recuperação de informação inestimável nos processos de organização do conhecimento. As designações terminológicas usadas para definir esta evolução não têm significado esclarecedor. Nesse caso, a computação como disciplina que deu uma funcionalidade diferente ao tesauro impôs sua terminologia ... A Web Semântica é uma oportunidade de usar thesauri nos processos de descrição e recuperação de informação. Para ilustrar isso, existem várias iniciativas de representação de tesauros usando RDF (Resource Description Framework) / XML (Extensible Markup Language), que levaram ao desenvolvimento do SKOS (Simple Knowledge Information System), um modelo de esquemas conceituais aplicáveis na Web Semântica e que está em desenvolvimento pelo World Wide Web Consortium (W3C).[33]
2014	MACULAN; LIMA	A despeito desse sistema conceitual expresso pelos diferentes tipos de relacionamentos, nem sempre é possível, mesmo aos humanos, apreender a especificidade do tipo de relacionamento, sobretudo dos relacionamentos associativos. Isso acontece porque esses relacionamentos não são explicitados de maneira mais formalizada, exigindo do usuário certo conhecimento do domínio representado. Em ambiente digital, contudo, essa característica é essencial para que possam ser inteligíveis à máquina.

[33] El tesauro ha ido evolucionando como una herramienta de recuperación de información inestimable en los procesos de organización del conocimiento.
Las designaciones terminológicas usadas para definir esa evolución no tienen un significado clarificador. En este caso la informática como disciplina que ha otorgado una funcionalidad diferente al tesauro ha impuesto su terminología...;
La Web semántica supone una oportunidad para utilizar los tesauros en los procesos de descripción y recuperación de información Para muestra de ello existen múltiples iniciativas de representación de tesauros mediante RDF (Resource Description Framework)/XML (Extensible Markup Language), que han dado lugar al desarrollo de SKOS (Simple Knowledge Information System), un modelo para esquemas conceptuales aplicables en la Web semántica y que actualmente se encuentra en desarrollo por el World Wide Web Consortium (W3C).

2014	LIMA; KROEFF; RIBEIRO JUNIOR	é uma linguagem documentária formada por descritores semanticamente relacionados e tem sido utilizado em diversos domínios, tais como ciência da informação, mecatrônica, folclore, jurídico, gênero e mulheres, dentre outras áreas do conhecimento. é um instrumento de controle terminológico e sua estrutura possibilita processo de tradução
2014	PIETERSE; KOURIE	aprimoramento oferecido pelos tesauros é a inclusão de relações semânticas além das relações hierárquicas. Essas relações têm como objetivo contribuir para o poder de raciocínio que deve ser construído em aplicativos que usam tesauros; é um termo técnico que se refere a um instrumento utilizado para controle de vocabulário, a fim de obter consistência na descrição dos itens e facilitar a recuperação. Um tesauro, nesse sentido, não define ou explica os conceitos do domínio, mas antes se concentra na descrição das relações entre os conceitos e usa notas de escopo principalmente para delinear o significado dos conceitos e resolver ambiguidades. Usamos o termo tesauro de recuperação de informação (tesauro de IR) para especificar isso. Por último, no campo IA, o termo se refere ao armazenamento de dados de palavras e frases que é comumente usado como um recurso de processamento de linguagem natural, daí o termo tesauro de processamento de linguagem natural (tesauro de PNL).[34]

[34] enhancement offered by thesauri is the inclusion of semantic relationships beyond hierarchical relationships. These relationships are intended to contribute to the reasoning power that is to be built into applications that use thesauri; is a technical term that refers to an instrument that is used for vocabulary control in order to achieve consistency in the description of items and to facilitate retrieval. A thesaurus in this sense does not define or explain the concepts of the domain, but rather focuses on describing the relations between the concepts and uses scope notes mainly to delineate the meaning of concepts and to resolve ambiguities. We use the term information retrieval thesaurus (IR thesaurus) to specify this. Lastly, in the AI field the term refers to the data store of words and phrases that is commonly used as a natural language processing resource, hence the term natural language processing thesaurus (NLP thesaurus).

2014	YOULIN; NUNES; ZHONGHUA	como um método de organização e representação do conhecimento, usam linguagem padrão e termos específicos para caracterizar assuntos e domínios bem definidos. Em geral, um tesauro é uma ferramenta de referência que lista palavras agrupadas de acordo com a similaridade de significado (contendo sinônimos e às vezes antônimos).[35]
2015	CARIBÉ; OLIVEIRA.	é uma linguagem documentária utilizada para traduzir o conteúdo dos documentos que fazem uso da linguagem natural para a linguagem do sistema, por meio de atribuição de descritores.
2016	VALENCIA; VOGEL; SILVA	Tesauro é um instrumento de trabalho usado para adequar os termos aos autores, usuários e indexadores em uma linguagem única. O Tesauro contribui com a organização da informação por meio da combinação de conceitos no por meio da indexação, visando à busca ou recuperação da informação (linguagem pós-coordenada). é um vocabulário controlado e dinâmico devido suas relações semânticas, sendo possível realizar relações entre dois termos de natureza diferentes
2017	SANTOS; LIMA; ARAÚJO	a estrutura dos tesauros tem sido adaptada conforme o aperfeiçoamento proporcionado pelas tecnologias de informação. Assim, as relações, associações e equivalências entre os termos possibilitam o cruzamento morfológico e semântico dos descritores que compõem os tesauros online, na medida em que auxiliam no trabalho de indexação e no momento da recuperação da informação. há técnicas e ferramentas de mineração de textos em estudos sobre inteligência artificial e linguística computacional, que vêm utilizando os tesauros como importantes instrumentos para a extração automática de informação, partindo de conjuntos de textos que resultam em uma série de aplicações, tais como a

[35] Thesauri, as a method of organizing and representing knowledge, use standard language and specific terms to characterize well defined subjects and domains. In general, a thesaurus is a reference tool that lists words grouped together according to similarity of meaning (containing synonyms and sometimes antonyms)

		indexação automática, a tradução automática, a sumarização de textos etc.
2017	BERTIR JUNIOR; SOERGEL; LIMA; MACULAN	Os tesauros são linguagens terminológicas que representam o conhecimento de determinado campo de especialidade, elaborados com base em um conjunto de regras, e são constituídos por um conjunto de termos descritores.
2017	SABBAG; GALDINO	Os tesauros são uma das linguagens documentárias mais utilizadas e que sofreram maior impacto das TIC's, tornando-se indispensável para uma indexação efetiva.
2017	MAIA; VASCONCELOS SOBRINHO; CUNDURÚ	um instrumento de organização, padronização e sistematização da informação. além de contribuir como ferramenta de auxílio ao profissional da informação na representação e recuperação da informação para tomada de decisões, auxilia, também, o usuário ao acesso da informação ambiental e o gestor na tomada de decisão
2017	GABRIEL JUNIOR; LAIPELT	Os tesauros são sistemas valiosos de Organização do Conhecimento (OC) que suportam a recuperação avançada de informações. A Web Semântica trouxe um interesse renovado nos tesauros como suporte para buscas semânticas e outros serviços de valor agregado. o que se pretende destacar no Thesa é a possibilidade de descrever as relações específicas, viabilizando desta forma o compartilhando de estrutura e dados em sistemas complexos, como o Linked data, e ainda manter a compatibilidade com modelos simplificados como da ISO e NISO... o Thesa vem a ser uma opção o desenvolvimento de vocabulários e tesauros, sendo um facilitador para o reuso de informação, alinhado a proposta de Dados Lincados do W3C e a Web Semântica

2017	FUJITA; ANDRADE E CRUZ; PATRÍCIO	é um sistema de organização do conhecimento adotado como ferramenta de representação documentária na indexação e na recuperação de informações cujo uso é difundido mundialmente. foram aprimorados para suprir as exigências dos antigos e dos novos campos e subcampos da ciência, já que estes necessitavam de ferramentas de organização da informação mais elaboradoras, específicas e flexíveis é uma linguagem de indexação de vocabulário controlado e pós-coordenado, o que oferece a essa ferramenta maior flexibilidade para atender às necessidades dos usuários, ao desenvolverem suas buscas.
2017	CAMPOS; CAMPOS; SOUZA	Linguagem artificial disseminada na área de representação e recuperação da informação são usados como instrumento para indexação e recuperação de informação em meio digital e outros instrumento para indexação e busca de conteúdos documentais instrumento de recuperação da informação
2017	ROSATI *et al*.	um recurso semântico sobre as características funcionais; contornam as questões de ambiguidade em linguagem natural, facilitando a identificação e integração das informações disponíveis em várias fontes de dados e permitindo que cientistas e aplicativos de computador interpretem mais efetivamente o significado dos dados.[36]

[36] a semantic resource on functional traits; bypass ambiguity issues in natural language, facilitating the identification and integration of the information available in multiple data sources and allowing both scientists and computer applications to interpret more effectively the meaning of data.

2018	MIRANDA; COELHO; DIAS.	É uma linguagem verbal (que usa termos da linguagem natural e que estão organizados alfabeticamente), pós-coordenada, pois permite a combinação dos termos no momento da indexação e no momento da busca e que, pode ser utilizada em ambientes automatizados. Essa combinação é feita por meio de artifícios, como os operadores booleanos (AND, OR, NOT).
2018	ORREGO	O Thesaurus é criado a nível local, sendo o conjunto de termos que permite uma gestão flexível no controle de vocabulários...[37]
2018	VÍQUEZ	O thesaurus integra sistemas de recuperação de informação de forma automatizada, aumenta a cooperação internacional, garante a troca de informação e o crescimento inteligente das redes de comunicação de dados como a Internet e a web. A plataforma tecnológica onde o tesauro é armazenado caracteriza-se por trabalhar com hipertexto nos descritores, o que permite a busca entre os mesmos descritores e vincular as buscas de descritores na web, permitindo ao pesquisador obter informações em diferentes formatos.[38]

[37] se crea Tesauro a nivel local, siendo el conjunto de términos que permite la gestión flexible en el control de los vocabulários.

[38] El tesauro integra sistemas de recuperación de información de forma automatizada, aumenta la cooperación internacional, garantiza el intercambio de información y el crecimiento inteligente de las redes de comunicación de datos como la Internet y la web;
La plataforma tecnológica donde se almacena el tesauro se caracteriza por trabajar con hipertexto en los descriptores, lo que permite realizar búsquedas entre los mismos descriptores y vincular las búsquedas del descriptor en la web, permitiéndole a quien investiga obtener información en diferentes formatos

2018	SANTOS; CERVANTES; FUJITA	e se constitui como instrumento apropriado para estabelecer a ligação entre o usuário e o conteúdo temático do acervo; tornou-se necessário para melhorar a busca e recuperação da informação em um contexto das TIC e proliferação, de forma exponencial, de informação por meio da internet. Esses recursos e acontecimentos muito contribuem para o aprimoramento de procedimentos de construção e disseminação de instrumentos como o tesauro.
2019	MORAIS; RAMALHO	A evolução dos tesauros é significativa e se faz presente nas normas que regulamentam sua utilização quando estas demonstram a necessidade de aplicação do controle de vocabulário em contexto digital. A Norma ISO 25964, publicada em 2011, aborda aspectos como interoperabilidade entre vocabulários controlados, formatos necessários para o intercâmbio de dados e a proposta de modelização de dados, denunciando a evolução desse tipo de linguagem documentária e seu aprimoramento quando da organização do conhecimento. Com o objetivo de favorecer uma maior aproximação entre as linguagens documentárias e as Tecnologias Semânticas, o consórcio W3C padronizou o Simple Knowledge Organization System – SKOS, um modelo de dados que possibilita a representação formal de tesauros em formatos legíveis por máquinas, favorecendo a interoperabilidade entre diferentes vocabulários.

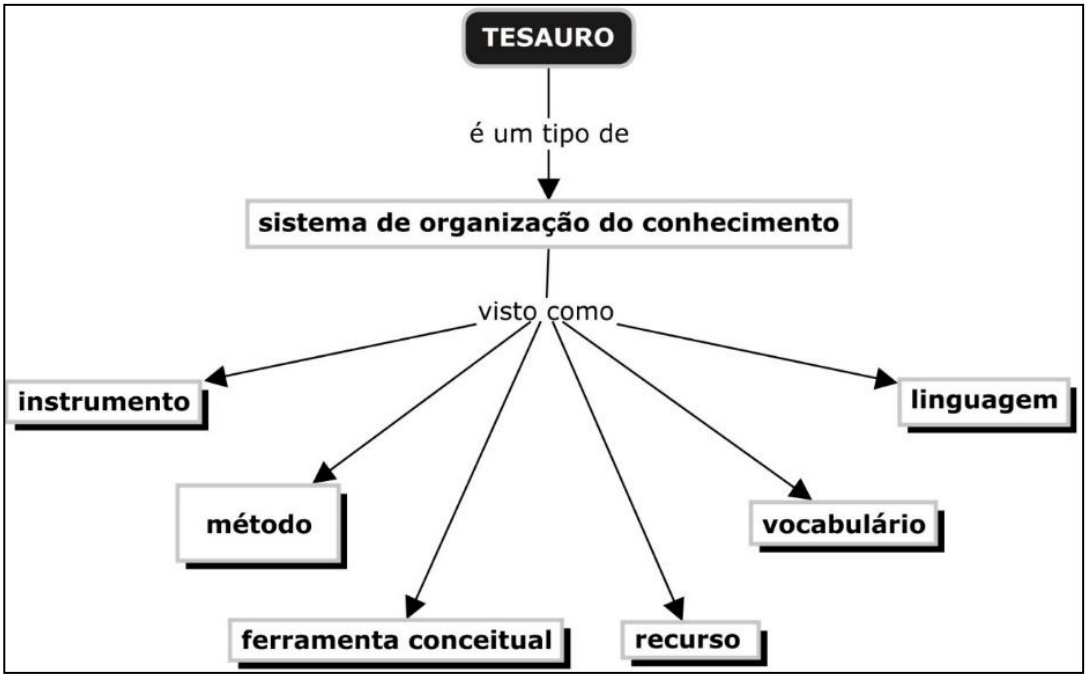
2019	VOGEL; KOBASHI	é um instrumento de controle do vocabulário que possibilita a indexação da informação, a recuperação desta com qualidade e a redução das ambiguidades oriundas da linguagem natural pode ser utilizado em várias etapas do gerenciamento de arquivos, como apoio à elaboração de tabelas de temporalidade, para estabelecer as formas de manuseio e armazenamento, e quem pode ter acesso aos documentos fato de poder ser aplicado a documentos e registros eletrônicos tanto quanto aos documentos em papel, por meio da construção de estruturas de diretórios nomeadas com as palavras-chave do tesauro

Fonte: Dados da pesquisa.

A riqueza conceitual sobre as concepções dos extratos acerca do sistema tesauro é inegável; tão como é inevitável não o perceber como um sistema que evoluiu como ferramenta para recuperação de informação nos processos de organização e representação da informação e do conhecimento (PEÑA; REYES; MOROY, 2013). Essa evolução é perceptível na escada conceitual representada no quadro acima. Antes visto como um sistema pré-coordenado, estruturado por listas alfabéticas estáticas, tesauro passou a ser um sistema pós-coordenado, apresentando-se como uma estrutura com maior flexibilidade para os profissionais da informação e seus usuários, bem como, consequentemente, para as atividades de organização, busca, representação e recuperação da informação.

Não se pretende, aqui, limitar um conceito como único e verdadeiro sobre as concepções acerca de tesauro, mas, sim, tem-se o intuito de esclarecer as distintas concepções sobre esse sistema. Para tanto, fora elaborado o mapa conceitual apresentado a seguir – figura 56, na tentativa de reunir os extratos identificados no quadro 7 apresentado anteriormente.

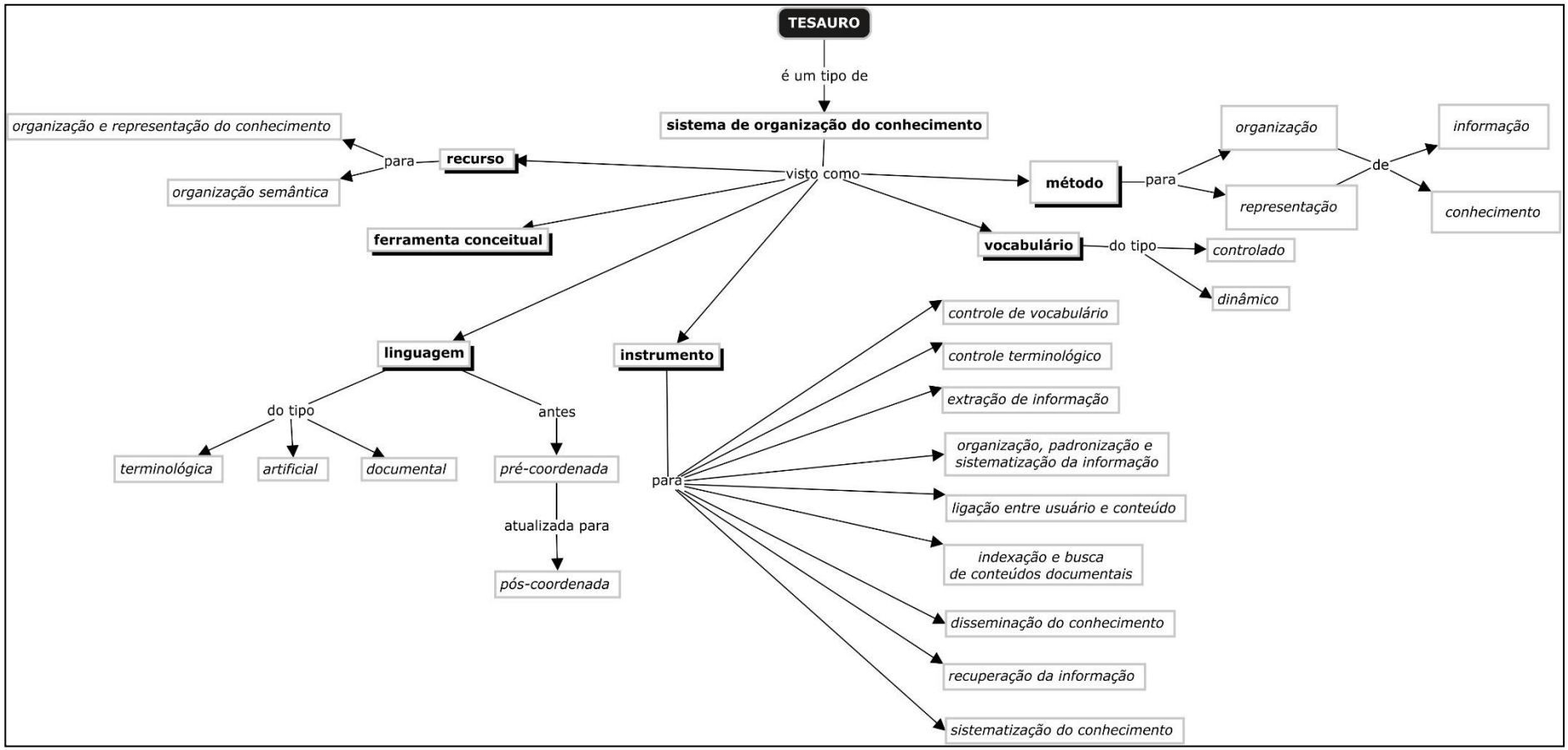
Figura 56 - Concepções sobre tesauros



Fonte: dados da pesquisa.

Ao observar o mapa sobre tesauro, é possível identificar que se trata de um tipo de sistema de organização do conhecimento visto como: **ferramenta conceitual; instrumento** para: controle de vocabulário; controle terminológico; extração da informação; organização; padronização e sistematização da informação; ligação entre usuário e conteúdo; indexação e busca de conteúdos documentais; disseminação do conhecimento; recuperação da informação; e sistematização do conhecimento; **linguagem** pós-coordenada, de caráter terminológico, artificial e documental; **método** para organização e representação da informação e do conhecimento; **recurso** semântico para organização e representação da informação e do conhecimento; **vocabulário** do tipo controlado e dinâmico. A extensão dessas concepções conceituais sobre tesauro, conforme figura 57, está apresentada a seguir.

Figura 57 - Concepções estendidas sobre tesauros



Fonte: O autor (2022) baseado em dados da pesquisa.

Além das concepções conceituais, que por um lado convergem e, por outro, divergem, outras percepções podem ser vistas no quadro apresentado. Por exemplo, quanto ao uso de tesauros: eles podem ser aplicados para resolver diferentes problemáticas informacionais, em qualquer tipo de ambiente, em diversos campos científicos e tecnológicos, “[...] embora em geral cumpram a função de homogeneizar terminologias, controlar vocabulários ou ajudar o usuário a navegar por um espaço de conceitos” (NAUMIS PEÑA; MORAN REYES; MORALES MONROY, 2013, p. 111).

Yulin, Nunes e Zhonghua (2014) enfatizam a ampla utilização do tesauro e suas aplicações de sucesso, mas frisam que esse sistema ainda apresenta pontos fracos, que estão cada vez mais aparentes, em razão do desenvolvimento acelerado da tecnologia da informação. Como exemplo: *ciclos longos de revisão do sistema*, *alto custo empregado nos processos de atualização*, *atraso severo de neologismo constantemente emergentes* e a *ausência da capacidade da atualização automática de tesauros*. Infere-se, portanto, que existe “[...] uma necessidade clara de melhorar as capacidades dos tesauros para aproveitar ao máximo os recursos de tecnologia da informação hoje” (YULIN; NUNES; ZHONGHUA, 2014, p. 132).

Moreiro González (2011) ainda destaca outros pontos que precisam receber maior atenção quanto ao desenvolvimento e uso de tesauros. Para o autor, esse sistema apresenta dificuldades na sua utilização, em razão das ambiguidades terminológicas e das potencialidades de inovação, podendo somar a este escopo a disponibilização crescente de documentos inseridos e compartilhados na web e a variedade de seus suportes. Além disso, dois outros pontos observados pelo autor se devem às “[...] dificuldades para descrever campos com alto nível de abstração (o da informática, por exemplo) ou documentos que não têm nenhum tipo de estrutura (texto de livre formato)” e ao “[...] custo elevado na criação, manutenção e funcionamento, evitáveis só com a automação ou semiautomação”. (MOREIRO GONZÁLEZ, 2011, p. 69).

Em razão dessas reflexões, apontadas como pontos a serem melhorados no desenvolvimento e uso de tesauros, em razão do aperfeiçoamento tecnológico, Naumis Peña, Moran Reyes e Morales Moroy (2013, p. 116) esclarecem que as tecnologias da informação utilizadas, para elaborar os conteúdos disponibilizados na web, necessitam de estruturas que permitam uma representação adequada dos

aspectos semânticos, isso porque “[...] os usuários requerem ferramentas conceituais e semânticas para uma organização eficaz da informação digital”, pois, conforme Moreiro González (2011, p. 71), “[...] os tesauros proporcionam respostas às perguntas dos usuários – o que continua sendo atual *e* o como supõe adaptar-se às novas necessidades”.

Quanto à estrutura do tesauro, fora apresentado que esse sistema é baseado em relacionamentos dos tipos: equivalência, partitivo, hierárquico e associativo. Poderia, aqui, considerá-lo como um sistema rico em relações, o que, de fato, é, mas essas relações, consideradas básicas,

> [...] ainda podem ser consideradas insuficientes para representar o conhecimento de um determinado domínio, uma vez que não anunciam, explicitamente, o significado dos relacionamentos, requisito que consideramos importante para a recuperação de informações em ambiente web. (MACULAN; LIMA; PIEROZZI JUNIOR, 2015, p. 4)

O estudo desenvolvido por Santos, Cervantes e Fujita (2018, p. 1116) corrobora com as reflexões acima, considerando que o sistema se tornou necessário para melhorar a busca e recuperação da informação, em um contexto de tecnologia da informação, comunicação e proliferação de informação no contexto web; e que o processo histórico e os recursos existentes são fundamentais para o aprimoramento de procedimentos de construção de tesauros.

Profissionais e pesquisadores da Ciência da Informação têm aplicado esforços para resolução dos problemas encontrados para desenvolvimento de SOC, inclusive de tesauros. Os pontos evidenciados acima, ao mesmo tempo que caracterizam os problemas advindos de tesauro, motivam pesquisas e aplicações desse sistema, tal como encorajam a manutenção de tesauros existentes para versões cada vez mais robustas. Por exemplo, ao tratar sobre o desenvolvimento de tesauros por automações ou semiautomação, pode-se citar dois projetos: Berti Junior, Soergel, Lima e Maculan (2017) em *Semiautomatização de relações em tesauros: uma proposta para refinamento de relacionamentos semânticos a partir do tesauro AGROVOC*, no qual visaram um modelo semiautomatizado para refinamento de relacionamentos semânticos em tesauros; e Santos, Moreira e Fujita (2018), ao apresentarem o artigo intitulado *Tesauro UNESP: integração do registro de autoridade para o ThemaTres*, cujo objetivo foi propor uma metodologia de integração semiautomática. Essas pesquisas evidenciaram o forte contributo

tecnológico para aprimorar o desenvolvimento de tesauros – o que corrobora a ideia de que o aperfeiçoamento das tecnologias de informação e comunicação aprimorou conceitos e processos de desenvolvimento de tesauros.

Quando Moreiro González (2011) citou os dois problemas percebidos, em razão do desenvolvimento de tesauro, ele também apresentou soluções para resolução desses, apontando que essas devem buscar reforço em outras ciências, estabelecendo um diálogo intercessor contínuo com outros campos do conhecimento científico e tecnológico. Aqui, portanto, aparecem a Linguística – no sentido de ampliar esse sistema, criando diferentes visões, em nível de indexação, de organização e representação da informação e do conhecimento; a Estatística – como "[...] solução matemática e, por isso, favorável a automação, no entanto distante da realidade linguística, pois os algoritmos dão uma visão simplista do mundo real"; e a Computação, que "[...] desde os anos 90, começaram a se aplicar modelos próprios de engenharia de software aos tesauros" (MOREIRO GONZÁLEZ, 2011, p. 70).

> A convergência destes métodos desembocou numa evolução dos standards sobre tesauros para verdadeiros mapas de representação do conhecimento. Esta transformação leva implícita a criação de relações adaptáveis a cada domínio, assim como a inclusão de grande variedade de descritores cuja categoria gramatical contribui com matizes na semântica do mapa conceitual e de novas categorias relacionais que ligam os recursos de informação e as facetas. (MOREIRO GONZÁLEZ, 2011, p. 70)

Ao tratar da criação de relações adaptáveis, percebe-se que "flexibilidade" foi um termo em evidência, quando mapeou-se as concepções conceituais sobre tesauros. Isso deve-se ao relacionamento entre semântica (outro termo evidente) e tecnologia. Peña, Reyes e Monroy (2013) destacam essa relação, principalmente em representações do espaço semântico.

> O espaço semântico é o meio em que estão envolvidas pelo menos duas pessoas, que podem ter diferentes percepções de língua, cultura, posição política, relação familiar, nível de escolaridade, entre outros, por isso é um espaço de troca de valores e significados, que afetam os comportamentos informativos. [...] a inferência em um sistema e em um espaço semântico consistiria em vincular a informação latente na rede e nas bases de dados com as necessidades de informação do usuário, a fim de estabelecer a maior precisão e relevância na busca realizada. (PEÑA; REYES; MONROY, 2013, p. 111)

Ainda de acordo com Peña, Reyes e Monroy (2013), a aparição do tesauro, enquanto sistema de organização do conhecimento, tratava de uma semântica simples, para representar uma classificação de seus descritores, isso por relacioná-los apenas de uma forma hierárquica a partir de generalização e especialização. Tão logo pode-se evidenciar, nesta tese, muitas iniciativas de profissionais e pesquisadores de Ciência da Informação para representação de tesauros, que levaram ao desenvolvimento do *Simple Knowledge Organization System* (SKOS). As características de SKOS estão apresentadas a seguir na figura 58.

Figura 58 - Características de SKOS

Fonte: O autor (2022) com base em (PEÑA; REYES; MONROY, 2013, p. 118).

O SKOS compreende sete propriedades, e, a cada propriedade, são agrupados rótulos (ou prefixos). Esses rótulos representam um conceito por uma expressão de linguagem natural; aqui: “[...] um conceito pode ser denotado por vários rótulos (dando expressão a sinônimos), enquanto um rótulo pode significar vários conceitos (dando expressão a homônimos)” (PEÑA; REYES; MONROY, 2013, p. 119).

As sete propriedades de SKOS se baseiam em: conceitos e seus respectivos esquemas e coleções; rótulos lexicais; relações semânticas; notações e propriedades de documentação e de mapeamento. Conforme mencionado, para cada propriedade existe um prefixo correspondente. Essas propriedades, e seus respectivos prefixos, são apresentados no quadro abaixo.

Quadro 8 - Propriedades SKOS.

PROPRIEDADES	PREFIXO
Classe conceitual e esquemas conceituais	skos:Concept skos:ConceptScheme skos:inScheme skos:hasTopConcept
Rótulos Lexicais	skos:prefLabel skos:altLabel skos:hiddenLabel
Relações Semânticas	skos:semanticRelation skos:broaderTransitive skos:narrowTransitive
Propriedades de Documentação	skos:note skos:ScopeNote skos:historyNote skos:changeNote skos:definition skos:editorialNote skos:example
Coleções Conceituais	skos:Collection skos:OrderedCollection skos:member skos:memberList
Notações	skos:notation
Propriedades de Mapeamento	skos:mappingRelation skos:exactMatch skos:broadMatch skos:narrowerMatch skos:relatedMatch

Fonte: Adaptado de Peña, Reyes, Monroy (2013, p. 119).

Apesar de Felipe e Silva Neto (2019, p. 30) apresentarem que: “[...] estudos recentes mostram que a produção e a publicação de tesauros têm diminuído significativamente, e que são poucas as áreas que ainda mantêm esforços relacionados com a produção de tesauros”, as pesquisas realizadas durante o processo de escrita desta tese evidenciam que o desenvolvimento e uso de tesauros se mantém constante, seja no aparecimento de novos tesauros ou de pesquisa orientadas à manutenção e atualização desses sistemas. Muitos deles, inclusive, aportando-se das propriedades SKOS, como o estudo mais recente publicado na revista P2P & Inovação, por Gabriel Junior e Laipelt (2020), intitulado *Descrição das relações semânticas para aplicação em KOS: uso do tesauro semântico aplicado (THESA)*, em que os autores tiveram como objetivo o desenvolvimento de um modelo de metacategorias semânticas para aplicação em KOS.

> O Tesauro Semântico Aplicado (THESA) foi desenvolvido com o objetivo de disponibilizar um instrumento de elaboração de tesauros para os estudantes de graduação em Biblioteconomia da UFRGS utilizarem na disciplina de Sistema de Classificação III, tornando-se também um objeto de pesquisa. O objetivo do Thesa é reduzir o trabalho operacional e possibilitar maior atenção ao trabalho de desenvolvimento cognitivo e conceitual referente a modelagem do domínio e na atribuição das relações paradigmáticas e sintagmáticas. O Thesa é distribuído em licença Creative Commons (CCBY) no Github (RENEFGJr/Thesa), podendo ser distribuído e instanciado em servidores de outras instituições, o software é livre para utilização, sendo objeto para fins didáticos em disciplinas dos cursos de graduação, para pesquisa na pós-graduação ou para instrumentalização de bibliotecas e uso profissional. (GABRIEL JUNIOR; LAIPELT, 2020, p. 124)

Ainda sobre o THESA, relacionando-o às problemáticas apresentadas sobre o desenvolvimento e uso de tesauros, e corroborando com as características das eras da interoperabilidade e do *big data*, bem como para solidificar a ideia da aplicação de aprendizado de máquinas e mineração de textos para desenvolvimento deste sistema, Gabriel Junior e Laipelt (2020, p. 125-127) reforçam que

> Para tornar o Thesa uma ferramenta mais completa e integrada ao SRI estão sendo desenvolvidas algumas Interfaces de programação de aplicações (API) para identificação de candidatos a termos a partir da análise de textos completos de uma base de dados. Essas ferramentas de conversão devem atuar não somente como instrumentos do indexador, mas como instrumento que deva interagir com o usuário em seu processo de busca de informação, bem como ser um instrumento que analisa as interações e a linguagem do usuário, por meio de seus *logs* de busca e dos documentos que ele selecionou e visualizou, com bases nessas interações. A análise dessas interações, bem como a incorporação de novos vocabulários, microtesauros e de tesauros devem viabilizar um processo de reindexação dos documentos na própria base de dados. O Thesa foi projetado não somente como ferramenta para descrever relações

> específicas, mas viabilizando o compartilhando de estrutura e dados em sistemas complexos, como o *Linked data*, mas também mantendo a compatibilidade com modelos simplificados como da ISO e NISO (p. 125).
> O Thesa possibilita interoperabilidade, comunicação com outros programas, utilizando os *Application Programming Interface* (API) ou Interface de programação de aplicações, que por meio de um *Token* (dispositivo eletrônico gerador de senhas) possibilita acesso de outras aplicações sem a necessidade de intervenção humana. (GABRIEL JUNIOR; LAIPELT, 2020, p. 127)

Outrora, Gabriel Junior e Laipelt (2017, p. 1) evidenciaram soluções problemáticas em torno de tesauros, principalmente no que tange a ideia de relações semânticas. O intuito foi: "[...] buscar uma reflexão sobre a aplicabilidade de requisitos semânticos de contextualização de domínio na modelagem de sistemas de organização do conhecimento, principalmente nos tesauros" e, com isso, apresentaram o THESA.

Quadro 9 - Exemplos de relações semânticas dos tesauros e Thesa.

Tesauro				**Thesa**		
Termo	**Relação**	**Termo**		**Conceito**	**Propriedade**	**Conceito**
Biblioteca Escolar	TG	Biblioteca		Biblioteca Escolar	é uma tipologia de	Biblioteca
Biblioteca	TE	Biblioteca Escolar		Biblioteca	tem como tipo	Biblioteca Escolar
Biblioteca digital	TG	Biblioteca		Biblioteca digital	tipo de acervo da	Biblioteca
Biblioteca	TE	Biblioteca digital		Biblioteca	tem acervo do tipo	Biblioteca Digital
Biblioteca Escolar	TR	Centros de Multimeios		Biblioteca Escolar	tem como parte (todo / parte)	Centros de Multimeios
Centros de Multimeios	TR	Biblioteca Escolar		Centros de Multimeios	faz parte da (todo / parte)	Biblioteca Escolar

Fonte: Gabriel Junior e Laipelt (2017, p. 6).

Ainda sobre o aporte das propriedades SKOS, para aperfeiçoamentos de tesauro, cita-se outro estudo, intitulado: *ThesaurusAPI: una API para la manipulación de tesauros*, de autoria de Martínez-González e Muñoz Nieto, publicado no periódico Scire em 2012. Aqui, os autores apresentaram a API discutindo sobre funcionalidades, tratamento e uso de SKOS. Para Matínez-González e Muño Nieto (2012, p. 41 e 42, tradução livre):

> O uso de tesauros em formato eletrônico requer a disponibilidade de ferramentas para construir e editar tesauros, ou consultá-los e usá-los em vários aplicativos. Nesse contexto, há a necessidade, comum no processamento de computadores, de pacotes de *software* genéricos que facilitem seu trabalho para desenvolvedores de aplicativos. As interfaces de programação de aplicativos (APIs) facilitam essa tarefa. Uma API fornece uma interface de programação geral que pode ser reutilizada em vários

> aplicativos. Assim, uma API para manipular tesauros facilita a construção de ferramentas que manipulam tesauros (p. 41).[39]
> ThesaurusAPI é uma API genérica para tesauros desenvolvida na Universidade de Valladolid. Suas características mais notáveis são: é uma API genérica para tesauros, independente do formato de representação dos tesauros manipulados, é fornecida como um pacote de distribuição gratuita, desenvolvida em Java, e suporta uma ampla funcionalidade, que permite tanto a consulta quanto a edição de thesauri, bem como importação e exportação de thesauri externos. Por outro lado, é uma API desenhada para facilitar o desenvolvimento de aplicações que manipulam tesauros que se armazenam localmente, dos quais, portanto, você tem o controle e a possibilidade de modificá-los de acordo com suas próprias necessidades. O objetivo é permitir o desenvolvimento de diferentes aplicações de usuário, que acessam os mesmos tesauros, mas que independem de modificações nos níveis inferiores do software: mudanças de *warehouse*, mudanças de esquema, etc. (...) A funcionalidade inclui a edição de tesauros, a sua consulta e a importação e exportação. A gestão da integridade é outro aspecto que recebeu atenção especial. Finalmente, são discutidas as vantagens e desvantagens de ter escolhido um padrão da Web Semântica, SKOS / RDF, para representar os tesauros armazenados. Usar o SKOS envolve o uso de um *warehouse* RDF para armazenar tesauros (p. 42).[40]

Apontam-se, ainda, outras pesquisas nacionais, como: Berti Junior, Soergel, Lima e Maculan (2017), Morais, Sousa e Ramalho (2019) e Ramalho (2015) internacionais, por exemplo: Gonzales Aguilar (2013), Rosati *et al*., (2017) e Youlin, Nunes e Zhonghua, (2014), que se aportam de estruturas SKOS ao tratarem de tesauros.

A criação de tesauros, bem como o processo de atualização de tesauros existentes, na recuperação dos extratos analisados, incidiu no descobrimento de sistemas que tratam de múltiplos objetos sociais, culturais, étnicos etc. Tesauro,

---

[39] La utilización de los tesauros en formato electrónico requiere la disponibilidad de herramientas que permitan construir y editar tesauros, o consultarlos y utilizarlos desde aplicaciones diversas. En este contexto se sitúa la necesidad, habitual en los tratamientos informáticos, de paquetes de software genéricos que faciliten su labor a los desarrolladores de aplicaciones. Las Application Programming Interfaces (API) facilitan esta tarea. Una API proporciona una interfaz de programación general que puede ser reutilizada desde diversas aplicaciones. Así pues, una API para manipular tesauros facilita la construcción de herramientas que manipulan tesauros (p. 41).

[40] ThesaurusAPI es una API genérica para tesauros desarrollada en la Universidad de Valladolid. Sus características más destacables son: es una API genérica para tesauros, independiente del formato de representación de los tesauros manipulados, se proporciona como un paquete de libre distribución, desarrollado en Java, y soporta una amplia funcionalidad, que permite tanto la consulta como la edición de tesauros, así como la importación de tesauros externos y su exportación. Por otro lado, se trata de una API pensada para facilitar el desarrollo de aplicaciones que manipulan tesauros que se almacenan localmente, de los cuales, por tanto, se tiene control y capacidad para modificarlos según las propias necesidades. El objetivo es permitir el desarrollo de aplicaciones de usuario diferentes, que acceden a los mismos tesauros, pero que son independientes de modificaciones en los niveles inferiores del software: cambios de almacén, cambios en los esquemas, etc. (...) La funcionalidad abarca la edición de tesauros, su consulta y la importación y exportación. La gestión de la integridad es otro aspecto al que se dedicó especial atención. Por último, se comentan las ventajas e inconvenientes de haber elegido un estándar de la Web Semántica, SKOS/RDF, para representar los tesauros almacenados. El uso de SKOS conlleva la utilización de un almacén RDF para guardar los tesauros (p. 42).

portanto, além de ser um sistema plural, em concepções conceituais, é um sistema que evidencia a representação dos mais diversos domínios científicos, tecnológicos e culturais. Essas descobertas estão apresentadas na Figura 59 a seguir, entre eles, tesauros e microtesauros em desenvolvimento, já desenvolvidos e em processo de atualizações.

Figura 59 - Exemplos de tesauro

Fonte: O autor (2022) baseado em dados da pesquisa.

O desenvolvimento de sistemas de organização do conhecimento, como tesauro, exige que profissionais de informação acompanhem as tendências tecnológicas e desenvolvam sistemas suscetíveis de manutenção, pois exige-se, cada vez mais, que esses sistemas apresentem estruturas flexíveis para inserção, modificação e correção dos termos inseridos, bem como dos relacionamentos entre eles. Essas manutenções podem ocorrer de forma automática ou semiautomática, por meio da validação de profissionais de informação, ou ainda podem ocorrer de forma manual.

Os processos que originam tesauros são orientados por métodos distintos. Existem as normas reguladoras para guiar profissionais de informação para construção desse sistema; mas, ainda assim, existem as operações adicionadas nos processos adotados para construir tesauros. Como um dos objetivos específicos desta tese, apresenta-se, a seguir, o mapeamento dos processos para desenvolver tesauro, de acordo com os extratos das pesquisas.

O mapeamento dos processos que originaram os tesauros buscou apresentar cada ação desenvolvida pelos pesquisadores. O mapeamento não aponta o domínio para o qual o sistema foi desenvolvido. Contudo, a descrição dos processos, para o desenvolvimento dos tesauros, compreende, para além da identificação do domínio, a idealização do sistema, apontando, por exemplo: objetivo, características particulares, modos de uso etc. Esse procedimento, de não identificar o domínio no mapeamento em si, se justifica por

> [...] existir a necessidade da criação de modelos baseados em teorias independentes de domínio, ou seja, que tome por base teorias e métodos capazes de representar domínios a partir de abordagens que não são construídas para resolver determinada situação, mas que compreendam questões teóricas e metodológicas que possam ser utilizadas na construção de modelos independentes de problemas específicos. (CAMPOS; MEDEIROS, 2012, p. 98)

Dos extratos selecionados para mapeamento, encontram-se aqueles que se aportaram de guias de construção de tesauros já publicados, como a pesquisa de Santos, Lima e Araújo (2017), que adotou três publicações distintas para fundamentar e desenvolver o sistema, como: o percurso metodológico descrito; o tutorial elaborado por Campos, Gomes e Motta (2004)[41]; as orientações de Vogel (2015)[42], as quais auxiliaram os pesquisadores na decisão da forma de apresentação dos termos no sistema, e, por fim, a aplicabilidade metodológica de Cavati Sobrinho e Fujita (2015)[43], para compreensão da análise do domínio.

Aponta-se, ainda, a pesquisa publicada por Vogel e Kobashi (2019, p. 57), intitulada: *Tesauro funcional para organização de arquivos administrativos*, que se

---

[41] CAMPOS, M. L. A.; GOMES, H. E.; MOTTA, D. F. **Elaboração de tesauro documentário**: tutorial. Rio de Janeiro: Biti, 2004. Disponível em: http://www.conexaorio.com/biti/tesauro. Acesso em: 06 ago. 2015.

[42] VOGEL, M. J. M. **Planejamento de Tesauros na Prática**. [*S.l.*]: Extralibris, 2015. [Notas de aula].

[43] CAVATI SOBRINHO, H.; FUJITA, M. S. L. Metodologia de construção de tesauro no domínio da Economia. *In*: PINTO, V. B.; VIDOTTI, S. A. B. G.; CAVALCANTE, L. E. (org.). **Aplicabilidades metodológicas em Ciência da Informação**. Fortaleza: Edições UFC, 2015. p. 139-153.

aportou do *National Archives of Australia* – Arquivo Nacional da Austrália (NAA) (2003)[44] e Smit e Kobashi (2003)[45]:

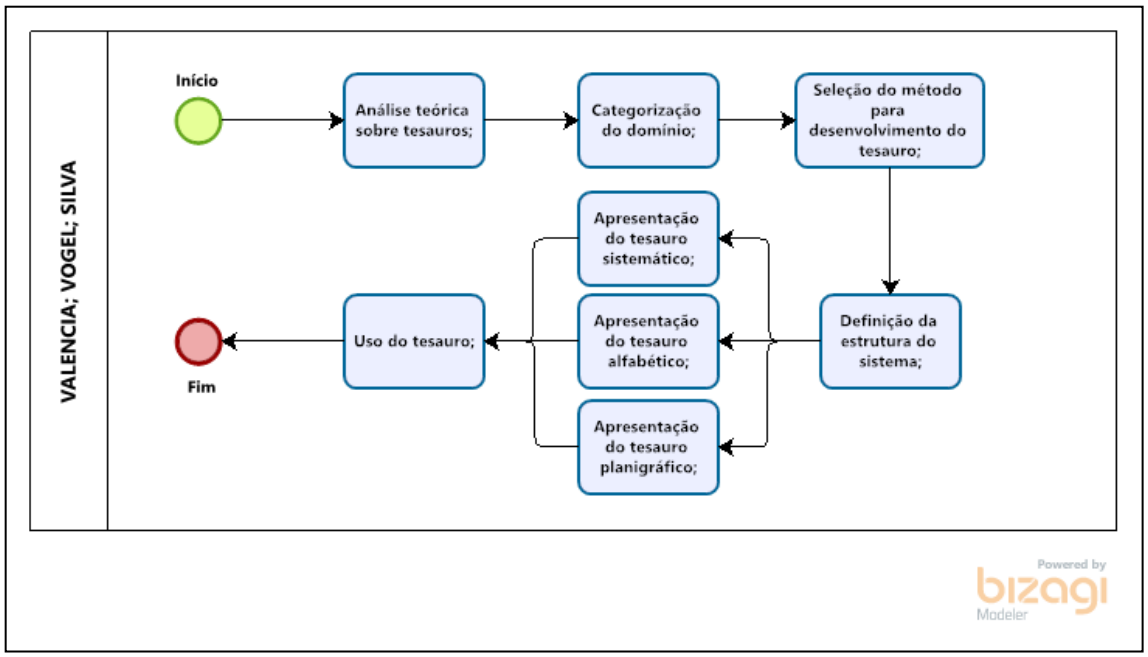
> O NAA traça recomendações organizadas em seis passos (AUSTRÁLIA. NSW State Archives and Records, 2003):
> ETAPA 1: Preparação. Antes de tudo é preciso saber se a instituição realmente precisa de um tesauro. Em caso positivo, nesta etapa, deve se definir quem fará o suporte gerencial do projeto, seu escopo e organizar os recursos necessários, humanos, financeiros e de programas de computador.
> ETAPA 2: Coleta de informação. É necessário identificar se existem procedimentos ou normas para atribuir títulos e classes, chamados pelo NAA de requerimentos para organização de arquivos, bem como conhecer o plano de classificação vigente. Além disso, é importante entrevistar usuários a fim de identificar suas necessidades de informação.
> ETAPA 3: Análise. Trata-se de avaliar toda a informação coletada, com atenção ao escopo do tesauro, ao plano de classificação de negócios e decidir quanto níveis o tesauro terá.
> ETAPA 4: Agrupamento. Neste momento, sugere-se identificar, em primeiro lugar, quais seriam as palavras-chave a serem adotadas do *Keyword* AAA ou quais precisariam ser criadas, bem como a necessidade de incluir notas de escopo. Na sequência, verifica-se as possibilidades de Descritores de Atividade do *Keyword* AAA ou quais precisariam ser criados e o mesmo deve ser feito com os Descritores de Conteúdo. Além disso, separam-se os termos preferidos dos não preferidos e, finalmente, atribuem-se os termos relacionados (RT ou TR).
> ETAPA 5: Busca de feedback. É o momento de testar o tesauro com seus usuários, buscando-se aprimorar os termos e suas relações. Verifica-se se o tesauro é consultado facilmente, se as notas de escopo fazem sentido, se há lacunas ou sobreposições, a fim de eliminar ambiguidades.
> ETAPA 6: Produção. Após os ajustes do feedback, o tesauro está pronto para ser publicado, seja no programa de computador escolhido, seja em livro ou página da internet. Finalmente, são estabelecidos os procedimentos para manutenção e revisão do tesauro. Destaque para a necessidade de documentar todas as decisões tomadas de modo a facilitar as modificações futuras.
>
> [...]
>
> Para a construção do plano, sugere-se (SMIT; KOBASHI, 2003, p. 24-25):
> Levantamento de uma lista livre de atividades;
> Análise crítica da lista, observando-se casos de sinonímias e a normalização gramatical;
> Elaboração de uma lista de atividades desdobradas em rotinas e acrescidas de remissivas;
> Categorização das atividades;
>
> Análise das nomeações, elaborando-se notas explicativas, caso seja necessário;
> Inclusão de uma codificação;
> Teste, avaliação e ajuste.

[44] AUSTRÁLIA. NSW State Archives and Records. **Developing and implementing a keyword thesaurus**. Kingswood: NSW State Archives and Records, 2003. Disponível em: https://www.records.nsw.gov.au/recordkeeping/advice/keyword-thesaurus. Acesso em: 25 fev. 2019.
[45] SMIT, J. W.; KOBASHI, N. Y. **Como construir vocabulários controlados para aplicação em arquivos**. São Paulo: Associação dos Arquivistas de São Paulo, 2003. Disponível em: http://www.arqsp.org.br/arquivos/oficinas_colecao_como_fazer/cf10.pdf. Acesso em: 12 fev. 2019.

Para Vogel e Kobashi (2019, p. 58), seguindo os modelos apresentados acima, “[...] obtém-se um instrumento para organização e recuperação das informações arquivísticas, que represente de fato suas características mais expressivas, isto é, suas funções dentro da instituição que as gerou”.

Enquanto Santos, Lima e Araújo (2017) trataram de tesauro na perspectiva do domínio da Física, e Vogel e Kobashi (2019) trataram do sistema no domínio dos Arquivos, Valencia, Vogel e Silva (2016) trabalharam sobre o tesauro de acervo espírita. Para tanto, os autores seguiram o processo mapeado a seguir, ilustrado na Figura 60.

Figura 60 - Mapeamento dos processos para construção do tesauro por Valencia, Vogel e Silva



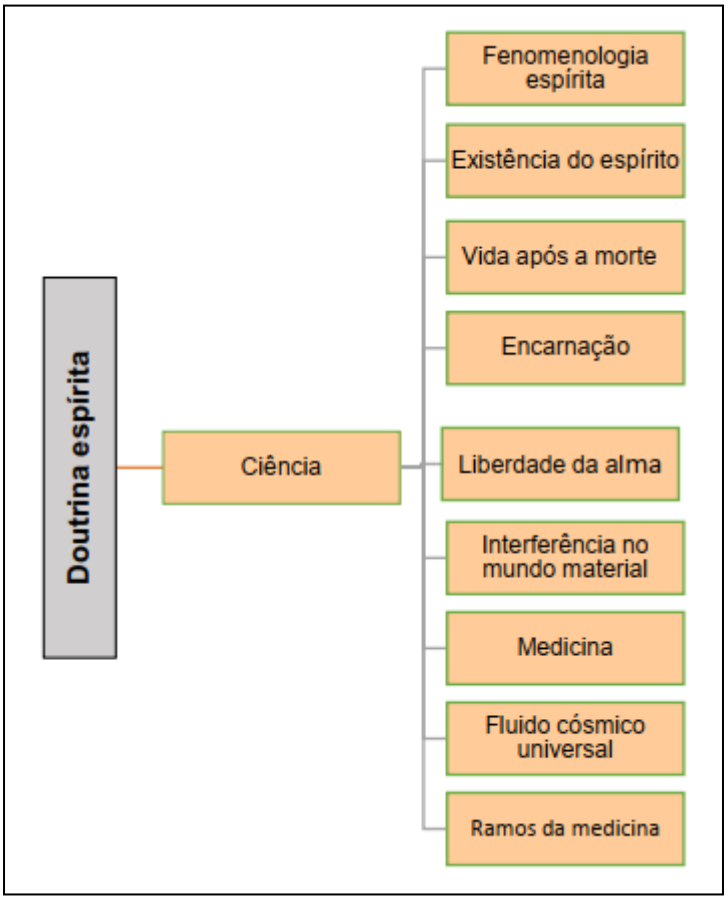
Fonte: O autor (2022) com base em Valencia, Vogel e Silva (2016).

A simplicidade na construção do tesauro sobre o acervo espírita se deve ao fato de esse sistema servir como um apoio de caráter estático, o qual “[...] apresenta os termos relacionados entre si, aplicados na indexação dos termos sobre o Espiritismo, facilitando a recuperação da informação sobre os assuntos relacionados ao tema” (VALENCIA; VOGEL; SILVA, 2016, p. 124).

Quanto à descrição das ações realizadas, primeiro iniciou-se como uma análise teórica sobre tesauro, seguida da categorização do domínio, através do estudo da CDD e da literatura especializada, como “*Vade Mecum* espírita, Dicionário de Filosofia Espírita, Glossários, Artigos espíritas, Vocabulário controlado, sites e

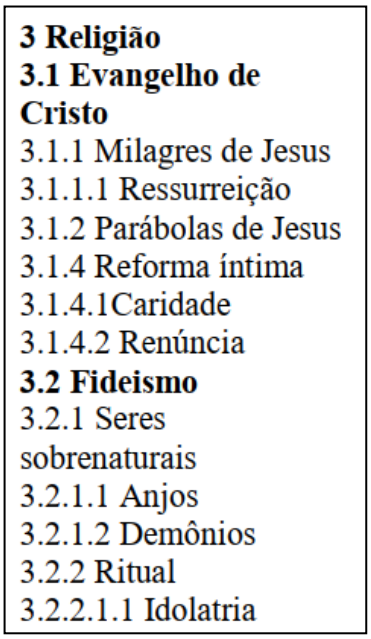
literatura da área” (VALÊNCIA; VOGEL; SILVA, 2016, p. 125); posteriormente foi selecionado o método misto de elaboração de tesauros – dedutivo e indutivo, o qual tratou da “terminologia utilizada nos materiais especializados, bem como a encontrada nos dicionários e nas enciclopédias” (VALÊNCIA; VOGEL; SILVA, 2016, p. 125); a etapa da definição da estrutura tomou norte do uso do tesauro, sendo apresentado em três formatos: planigráfico, sistemático e alfabético. O exemplo de cada um desses formatos pode ser visto a seguir em destaque nas Figuras 61, 62 e 63.

Figura 61 - Apresentação planigráfica



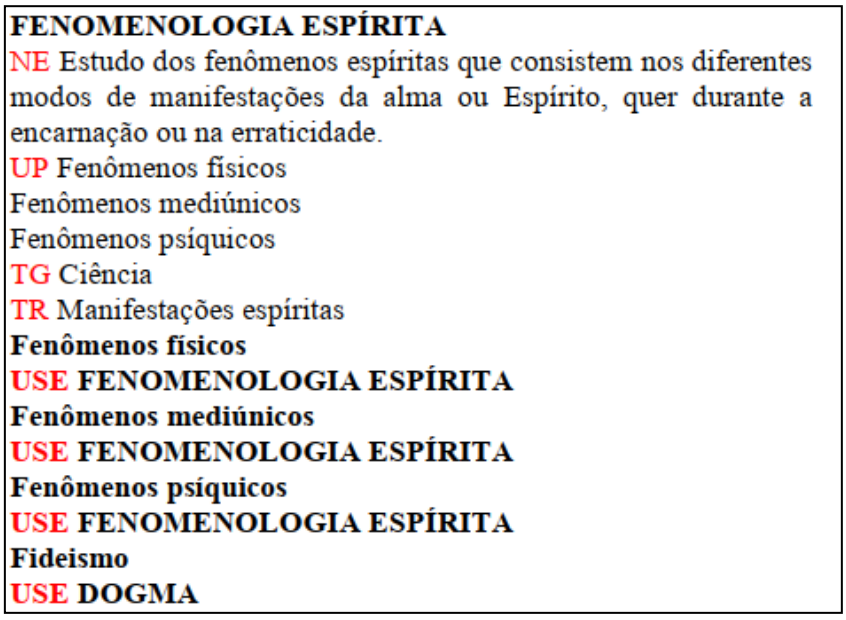
Fonte: Valencia, Vogel e Silva (2016, p. 134).

Figura 62 - Apresentação sistemática



Fonte: Valencia, Vogel e Silva (2016, p. 138).

Figura 63 - Apresentação alfabética



Fonte: Valencia, Vogel e Silva (2016, p. 140).

O tesauro proposto objetiva, também, auxiliar os usuários “[...] a recuperarem a informação necessária para a continuidade de seus estudos e leituras” (VALENCIA; VOGEL; SILVA, 2016, p. 130). Apesar de ele ser idealizado em um formato estático, as autoras perceberam a necessidade de desenvolver esse tesauro sob domínio das tecnologias digitais, por meio da utilização de softwares específicos para construção desse sistema. Essa necessidade é corroborada ao relatarem que “consequentemente propõe a

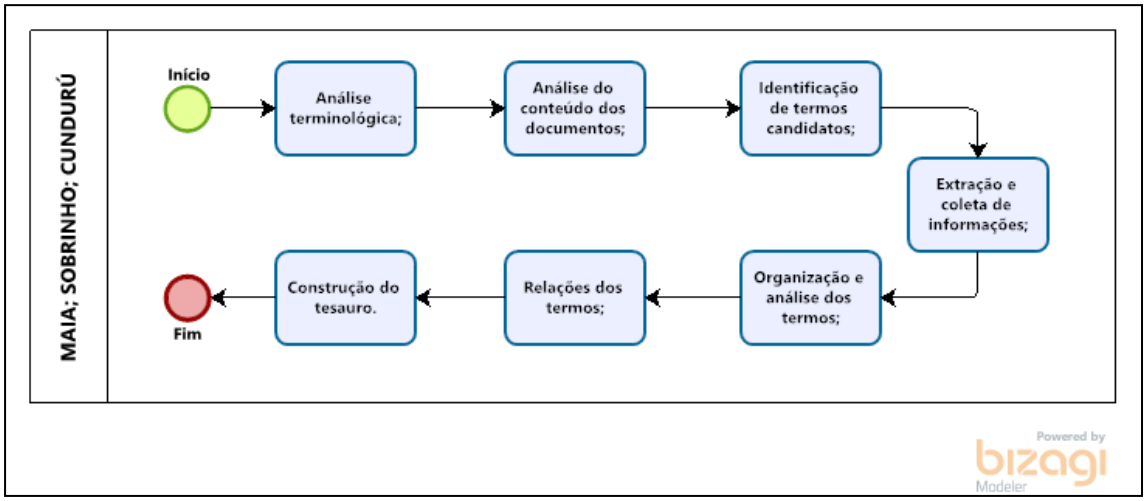
utilização de softwares, que possibilitem o desenvolvimento desse instrumento" e que, para isso, será necessário analisar o tesauro quanto a adequação "para a proposta do trabalho, prevalecendo a questão da compatibilidade, o compartilhamento desses dados e seu gerenciamento", analisando, também, a "[...] questão da compatibilidade das linguagens existentes, facilidade e uso e a questão de custo"(VALENCIA; VOGEL; SILVA, 2016, p. 130). Por fim, as autoras acreditam que "[...] será por meio da utilização efetiva do Tesauro nas bibliotecas especializas, que saberemos se a ferramenta atende efetivamente à indexação e recuperação da informação sobre o Espiritismo" (VALENCIA; VOGEL; SILVA, 2016, p. 131).

Um outro tesauro aqui tratado é a proposta da pesquisa *Terminologia aplicada à produção científica sobre gestão ambiental: diretrizes à elaboração de um microtesauro*, de autoria de Maia, Vasconcelos Sobrinho e Cundurú, publicada em 2017. A pesquisa

> [...] analisa as terminologias ambientais utilizadas na produção científica dos Programas de Pós-Graduação (PPGs) da Amazônia brasileira das áreas temáticas Interdisciplinar e Ciências Ambientais por meio da análise documental e de conteúdo das teses e dissertações. Traçam-se diretrizes para elaboração de um microtesauro terminológico no campo da gestão ambiental.
> As diretrizes para a criação do microtesauro como instrumento terminográfico apresentam conceitos a cada descritor ou termo coletado das terminologias em gestão ambiental identificadas na produção científica dos PPGs. Essas são apresentadas na forma da relação de conceitos como instrumento para normalizar a linguagem desta área específica, possibilitando, assim, o entendimento da sua estrutura conceitual. (MAIA; VASCONCELOS SOBRINHO; CUNDURÚ, 2017, p. 83)

O mapeamento do microtesauro construído pelos autores está apresentado na Figura 64.

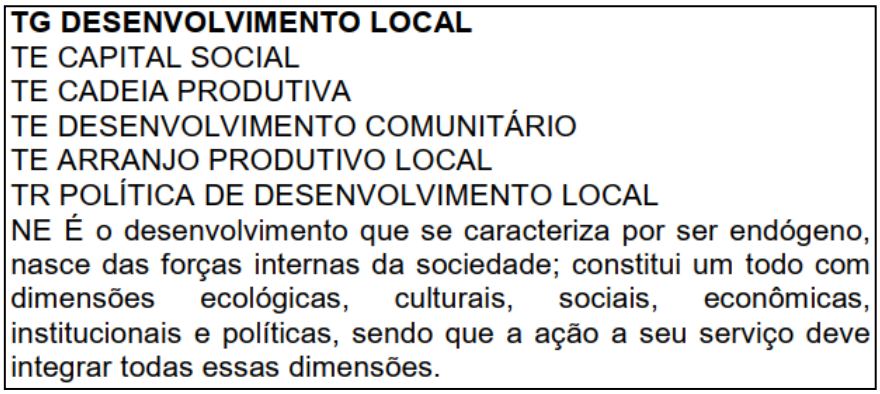
Figura 64 - Mapeamento dos processos para construção do tesauro por Maia, Sobrinho e Cundurú



Fonte: O autor (2022) com base em Maia, Vasconcelos Sobrinho e Cundurú (2017).

Iniciando o processo para construção do tesauro em questão, os autores realizaram uma análise da terminologia da área de gestão ambiental; tendo realizado, para tanto, “[...] um estudo exploratório no portal (sites) dos Programas de Pós-Graduação (PPG) credenciados pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) na Região Norte” do Brasil; a “[...] intenção foi coletar informações pertinentes ao desenvolvimento do estudo, conhecendo as áreas temáticas e linhas de pesquisa dos programas para posterior análise terminológica”. Tão logo “[...] o estudo seguiu uma abordagem qualitativa, com ênfase na análise do conteúdo dos documentos (os significados da terminologia ambiental)”; além de “[...] fazer as relações dos termos e conceitos ambientais para posterior elaboração de um microtesauro sobre gestão Ambiental” (MAIA; VASCONCELOS SOBRINHO; CUNDURÚ, 2017, p. 91). Para a etapa de extração e coleta dos termos e conceitos na área ambiental, os autores aportaram-se de um questionário como instrumento para a ação. O questionário “[...] teve uma função importante na coleta, extração, organização e análise das informações, pois, através de sua aplicação, foi possível realizar uma leitura técnica e aprofundada da produção científica”. Assim, o “[...] registro dos termos foi realizado após a leitura técnica e a identificação dos conceitos e termos no momento da análise do conteúdo dos documentos” (MAIA; VASCONCELOS SOBRINHO; CUNDURÚ, 2017, p. 92).

Figura 65 - Apresentação de parte das relações do microtesauro sobre gestão ambiental



Fonte: Adaptado de Maia, Vasconcelos Sobrinho e Cundurú (2014).

Para Maia, Vasconcelos Sobrinho e Cundurú (2017, p. 95)

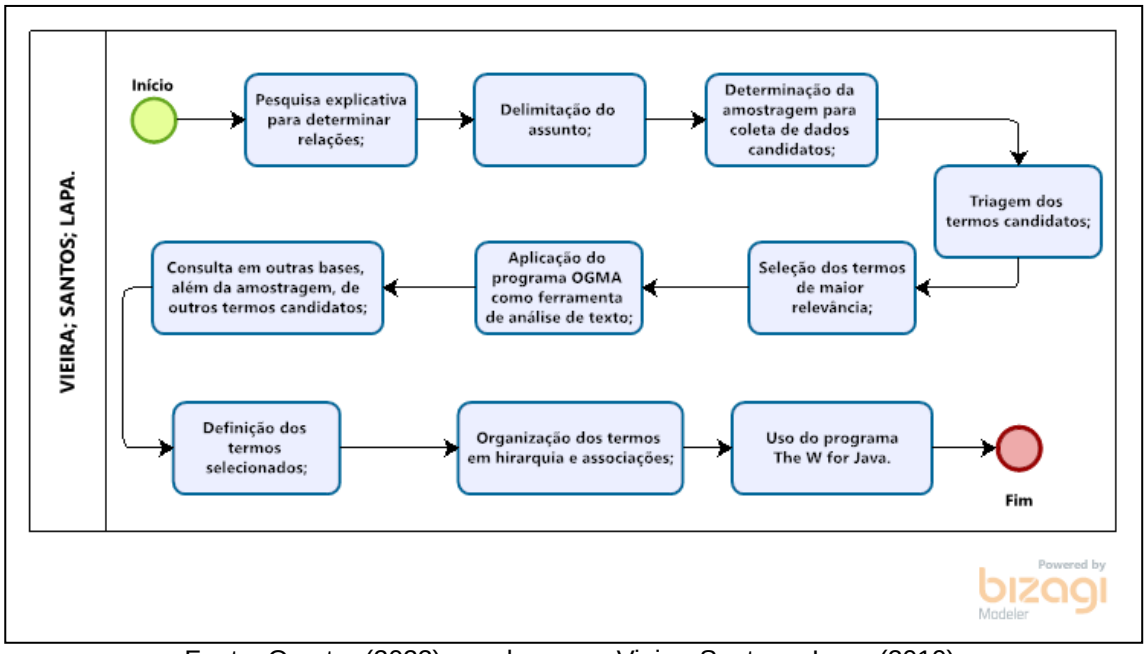
> A elaboração do microtesauro terminológico possibilitou identificar os termos próprios do campo da gestão ambiental, fornecer referências para a compreensão de termos e conceitos, organizar e divulgar informações científicas ambientais e constituir-se em um instrumento terminológico para representação e transmissão do conhecimento especializado na temática gestão ambiental, sendo uma importante fonte de informação em meio ambiente.

O mapeamento da pesquisa descrita acima pode ser visto como um tesauro comum, um tesauro estático, assim como o tesauro espírita. Talvez sejam esses tipos de tesauros que Felipe e Silva Neto (2019) tenham se referido, quando enfatizam que a produção e publicação de tesauros têm diminuído. Contudo, reforçar-se, mais uma vez, que as pesquisas realizadas para desenvolvimento desta tese orientam para um outro olhar, uma outra percepção; a de que a construção desses sistemas tem sido cada vez mais frequente, mas com outras formas e, por assim dizer, outras representações, as que vão além do suporte físico e migram para as instâncias tecnológicas digitais. No sentido de confirmar essa percepção, os mapeamentos a seguir corroboram a ideia de que o aprimoramento tecnológico digital aperfeiçoou os processos para construção de tesauros.

Vieira, Santos e Lapa (2010), no artigo: *Estudo da construção e aplicação do tesauro na recuperação da informação de teses e dissertações do Programa de Pós-Graduação em Desenvolvimento Urbano*, buscaram “[...] apresentar o processo de desenvolvimento e os principais resultados observados nas etapas de planejamento

e elaboração de um tesauro" (VIEIRA; SANTOS; LAPA, 2010, p. 71). A seguir, na Figura 66, está a apresentação do mapeamento dos processos.

Figura 66 - Mapeamento dos processos para construção do tesauro por Vieira, Santos e Lapa.



Fonte: O autor (2022) com base em Vieira, Santos e Lapa (2010).

A primeira ação buscou entender o tesauro e, para tanto, adotaram como técnica a "[...] revisão de literatura, análise e discussão, baseada em artigos, teses, dissertações e livros, que versem sobre SRI e Tesauros"; a delimitação do assunto foi realizada na área do conhecimento de Arquitetura e Urbanismo, em página consultada sob domínio da UFPE; na determinação da amostragem para coleta dos dados, adquiriu-se "[...] uma amostragem de cento e doze teses e dissertações, da área eleita para servir de modelo a construção do tesauro"; tão logo foi realizada a triagem dos termos "[...] através da ordem de frequência em que apareciam, selecionando assim, os termos compreendidos como de maior relevância para ser incorporados como linguagem controlada" e, para finalizar esta etapa, contou-se com a aplicação do programa OGMA[46], "[...] que funciona como ferramenta de análise de texto e permite a apuração da quantidade de ocorrência das palavras no texto". (VIEIRA; SANTOS; LAPA, 2010, p. 76).

Os autores ressaltam que "[...] outros termos foram incorporados à construção da linguagem controlada sendo extraídos de dicionários ou tesauros especializados na área de arquitetura"; e que, pela dificuldade de localizar todos os "[...] descritores

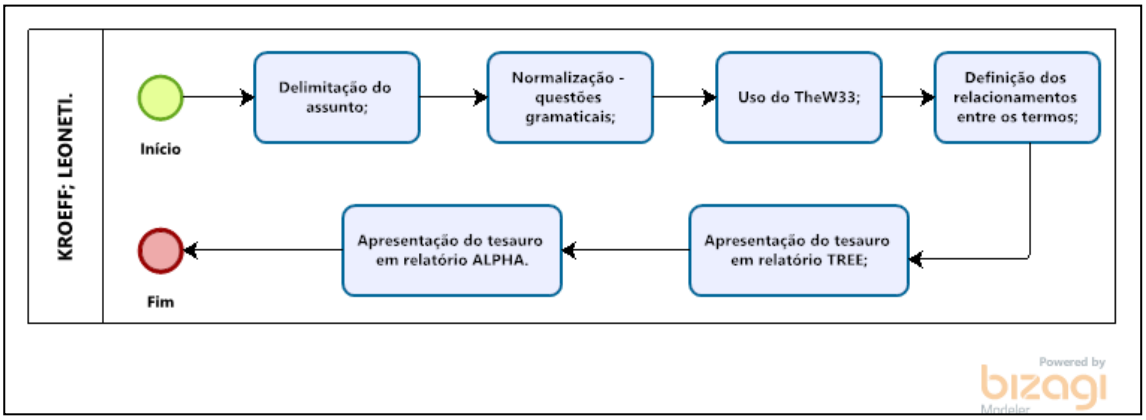
[46] https://www.luizmaia.com.br/ogma/

selecionados nos passos anteriores, alguns verbetes foram consultados em outras fontes de informação, como as enciclopédias virtuais, a exemplo da Wikipédia[47]", "[...] após a obtenção da definição de todos os termos selecionados como descritores, o procedimento seguinte foi o de organizá-los segundo sua hierarquia e associações"; para esta etapa, os autores contaram com o auxílio do programa The W forJava[48], software de código aberto "[...] que permite a elaboração de um tesauro virtual ao estruturar os termos que foram dispostos hierarquicamente de acordo com suas relações". (VIEIRA; SANTOS; LAPA, 2010, p. 76).

Ao seguir as etapas do mapeamento dos processos, para desenvolvimento do tesauro, além de chegarem ao sistema em si, os autores ainda perceberam outras formas de representação do domínio e originaram um glossário arquitetônico e um mapa conceitual, representando "[...] os termos estabelecidos na linguagem controlada, assim como suas relações" (VIEIRA; SANTOS; LAPA, 2010, p. 77).

Em 2012, Kroeff e Leoneti publicaram o *Estudo preliminar do tesauro brasileiro de Ciências do Esporte*, que apresentou um tesauro ao qual se referencia no título. Para desenvolvimento do tesauro, os autores seguiram as etapas apresentadas na Figura 67.

Figura 67 - Mapeamento dos processos para construção do tesauro por Kroeff e Leoneti



Fonte: O autor (2022) com base em Kroeff e Leoneti (2012).

A descrição dos processos inicia-se pela ação de delimitar o assunto onde os autores tiveram apoio em dois materiais bibliográficos, sendo eles o Dicionário de

[47] https://pt.wikipedia.org/wiki/Wikip%C3%A9dia:P%C3%A1gina_principal
[48] https://publish.uwo.ca/~craven/freeware.htm

educação física e esporte[49], de Barbanti, e o Dicionário enciclopédico Tubino do esporte[50]. Em sequência, seguiram para os procedimentos de análise e construção do sistema, seguindo: a normalização, as questões gramaticais, dispondo de palavras no singular e o uso de termos compostos; nesta etapa, "[...] foram adicionadas notas de escopo, aos termos específicos (*narrower term* – NT), a fim de definir as principais abordagens. Para as definições de notas de escopo utilizou-se os dicionários"; "[...] para a elaboração e consolidação do estudo preliminar de tesauro foi utilizado o software livre TheW33[51]". "O 'TheW33' é um programa para o *Microsoft Windows* e NT para criação e manutenção de Tesauros" (KROEFF; LEONETI, 2012, p. 83).

A etapa de relacionamento entre os termos apresentou os elementos de nota de escopo: use, use para/usado para: termo geral, termo específico e termo relacionado. Em seguida, o tesauro foi apresentado em dois formatos, a saber: ALFA: "[...] formato que lista todos os termos do tesauro de maneira alfabética, incluindo todos os relacionamentos e notas que foram feitos com o termo"; e TREE: "[...] formato que lista os termos do tesauro em formato de árvore hierárquica (notas de escopo não são exibidas nesse formato)" (KROEFF; LEONETI, 2012, p. 84).

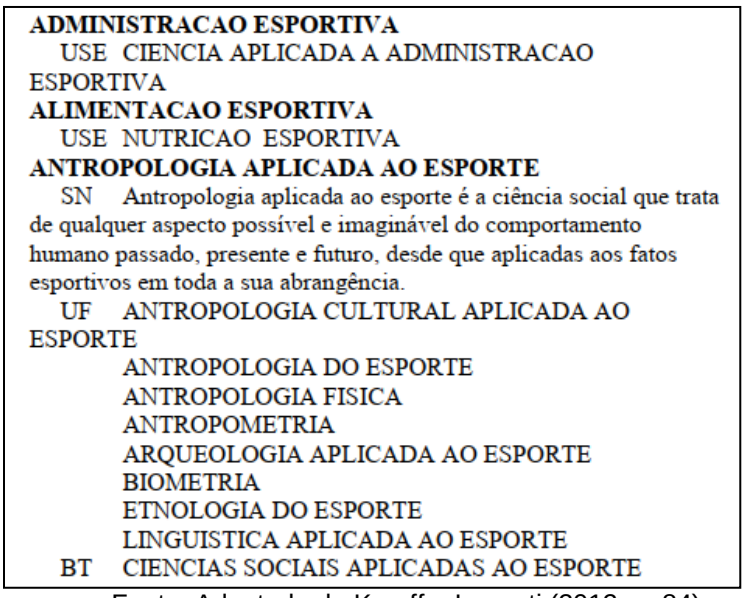
Figura 68 - Representação em formato Tree do tesauro brasileiro de ciências do esporte

CAMPOS CIENTIFICOS INTERDISCIPLINARES DO ESPORTE
NTCIENCIA DA MOTRICIDADE HUMANA
CIENCIA DO ESPORTE
CIENCIA POLITICA APLICADA AO ESPORTE
CIENCIAS APLICADAS AO TREINAMENTO ESPORTIVO
CIENCIAS APLICADAS AS ATIVIDADES FISICAS/ESPORTIVAS ADAPTADAS
ESPORTE COMPARADO
INFORMACAO E DOCUMENTACAO APLICADAS AO ESPORTE
CAMPOS CIENTIFICOS RELACIONADOS A AVALIACAO E AS MEDIDAS NO ESPORTE
NTCINEANTROPOMETRIA
DERMATOGLIFIA
SOMATOTIPOLOGIA

Fonte: Adaptado de Kroeff e Leoneti (2012, p. 84).

[49] BARBANTI, V. J. **Dicionário de educação física e esporte**. 2. ed. Barueri: Manole, 2003.
[50] TUBINO, M. J. G.; TUBINO, F. M.; GARRIDO, F. A. C. **Dicionário enciclopédico Tubino do esporte**. Rio de Janeiro: SENAC, 2007.
[51] https://publish.uwo.ca/~craven/freeware.htm

Figura 69 - Representação em formato Alpha do tesauro brasileiro de ciências do esporte



ADMINISTRACAO ESPORTIVA
USE CIENCIA APLICADA A ADMINISTRACAO ESPORTIVA
ALIMENTACAO ESPORTIVA
USE NUTRICAO ESPORTIVA
ANTROPOLOGIA APLICADA AO ESPORTE
SN Antropologia aplicada ao esporte é a ciência social que trata de qualquer aspecto possível e imaginável do comportamento humano passado, presente e futuro, desde que aplicadas aos fatos esportivos em toda a sua abrangência.
UF ANTROPOLOGIA CULTURAL APLICADA AO ESPORTE
ANTROPOLOGIA DO ESPORTE
ANTROPOLOGIA FISICA
ANTROPOMETRIA
ARQUEOLOGIA APLICADA AO ESPORTE
BIOMETRIA
ETNOLOGIA DO ESPORTE
LINGUISTICA APLICADA AO ESPORTE
BT CIENCIAS SOCIAIS APLICADAS AO ESPORTE

Fonte: Adaptado de Kroeff e Leoneti (2012, p. 84).

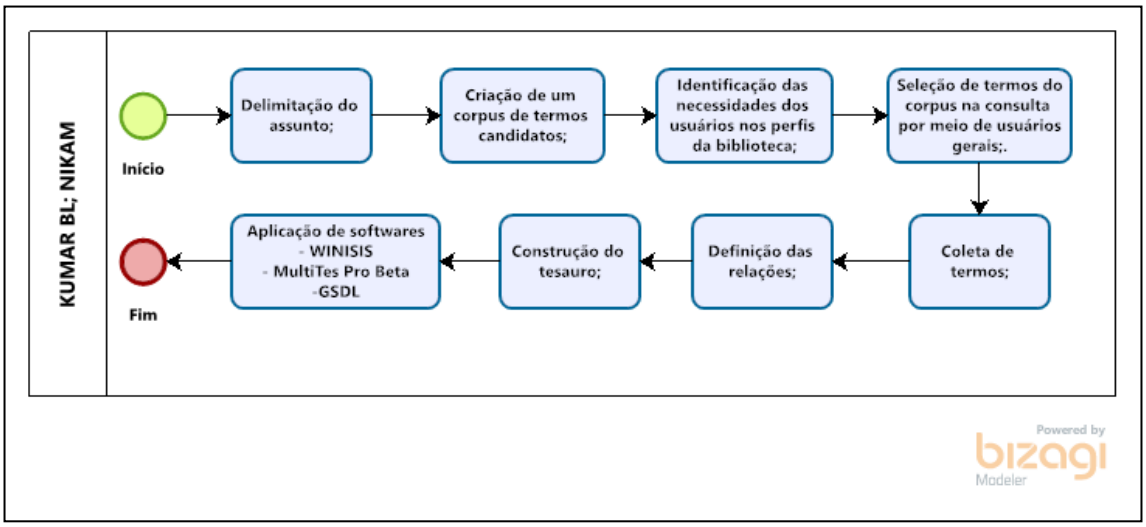
Um dos processos mais importantes para se executar, antes de dispor do uso do sistema, é a validação por parte de especialistas, etapa essa que, até agora, pode-se perceber que não foi apresentada nos processos mapeados até aqui. Essa importância é reconhecida pelos autores quando indicam a necessidade de que as classes principais e subclasses “[...] sejam revisadas por especialistas da área dos esportes e da educação física”, bem como que “[...] seja realizada a avaliação do sistema pelos pares profissionais da área de biblioteconomia” (KROEFF; LEONETI, 2012, p. 99).

A publicação da pesquisa descrita acima não apresenta o tesauro em seu formato final. Para tanto, os autores esclareceram que, após sua conclusão, se pretende divulgar o sistema aos autores das obras que deram origem ao tesauro e colocar o sistema à disposição. Mas, para a publicação do sistema final, será necessário passar pela etapa de validação e, assim, poderá ser “[...] disponibilizado eletronicamente em versão on-line (disponível na World Wide Web) para consulta aos usuários” (KROEFF; LEONETI, 2012, p. 99).

A ciência do esporte foi tema do tesauro mapeado e caracterizado acima. Outra ciência foi alvo de estudos, desta vez a ciência da ioga. O trabalho de Kumar Bl e Nikam (2014) buscou desenvolver um sistema de apoio à informação para a ciência da ioga, “[...] usando sistemas de organização do conhecimento

especialmente projetados, como um glossário de ioga e um dicionário de sinônimos de ioga"; na pesquisa, um "[...] tesauro para ciência da ioga é desenvolvido usando o software Greenstone Digital Library". O mapeamento dos processos deste tesauro está apresentado a seguir na Figura 70.

Figura 70 - Mapeamento dos processos para construção do tesauro por Kumar Bl e Nikam



Fonte: elaborado pelo autor com base em Kumar Bl e Nikam (2014).

Iniciando pela ação de delimitar o assunto, os autores buscaram estudar o domínio por meio da literatura disponível, como: dicionários, glossários, enciclopédias, artigos em periódicos, anais de seminários, conferências e simpósios. Em seguida, partiram para a fase de criação de um corpus de termos e para a identificação das necessidades dos usuários nos perfis da biblioteca de ioga. Essa última fase foi primordial para seguir na construção do sistema – a fase de seleção de termos foi feita por meio de estudos com especialistas no assunto: oficinas de pesquisas, cientistas, diretores, docentes, discentes; muitos dos envolvidos integram a universidade de ioga. Tão logo que realizado as ações para a seleção dos termos, foram assim definidas as devidas relações (KUMAR BL; NIKAM, 2014).

O registro dos termos foi evidenciado de acordo com as discussões estabelecidas pelos especialistas, elegendo aqueles mais convenientes. Aqui, pode ser definido o *layout* do tesauro na construção do sistema. Para desenvolvimento do tesauro, os autores, inicialmente, utilizaram o *software* WINISIS[52]; o MultiTes Pro

[52] http://www.unesco.org/new/en/communication-and-information/information-society/open-source-and-low-cost-technologies/information-processing-tools/cdsisis-database-software/cdsisis-for-window/

Beta[53], para criação de um glossário bilíngue interligado ao tesauro. O uso desse *software* se deveu pela importação dos dados, facilitando a criação de campos personalizados e a geração de relatórios alfabéticos, que puderam ser carregados em ambiente *web*. Contudo, mais tarde, os autores optaram por substituir o MultiTes pelo software Greenstone Digital Library Software (GSDL), na justificativa do anterior não ser um programa de código aberto e nem um *freeware*[54]. Os autores justificam o uso do GSDL por permitir hiperlinks interna e externamente. (KUMAR BL; NIKAM, 2014).

A Figura 71 representa o produto gerado pela pesquisa, um dicionário de sinônimos de ioga em sânscrito-inglês.

Figura 71 - Dicionário de sinônimos de ioga com hiperlink de glossário de ioga em sânscrito-inglês

Fonte: Kumar Bl e Nikam (2014, p. 292).

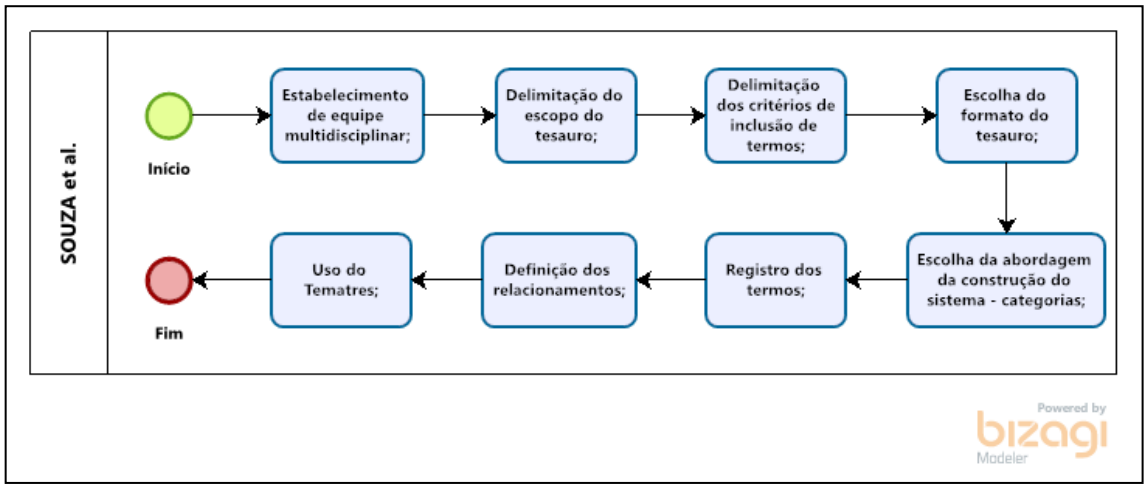
As ações de consulta aos especialistas devem ser empreendidas como uma ação constante, ou seja, do momento em que se planeja o sistema, a idealização dos processos, o fluxo de atividades, as determinações dos termos e os relacionamentos, a aplicação dos termos em sistemas de informação, as

[53] http://www.multites.com/
[54] Software gratuito.

funcionalidades determinadas para uso, a validação dessas funcionalidades e uso do tesauro. Isso inclui o diálogo interdisciplinar com os mais variados profissionais, muitas vezes pela mobilização de uma unidade ou central de informação em sua totalidade. Exemplo disso foi o desenvolvimento do Tesauro do Regime Militar Brasileiro, em que a pesquisa visou a construção do "projeto conjunto do CPDOC (Centro de Pesquisa e Documentação de História Contemporânea do Brasil) e da Escola de Matemática Aplicada", duas unidades pertencentes à Fundação Getúlio Vargas; o trabalho teve o "[...] objetivo de construção de tesauros retratando período da história contemporânea no Brasil"; apresentando, assim, um "[...] tesauro que retrata o período do Regime Militar Brasileiro – também conhecido como período de ditadura militar" (SOUZA *et al*., 2014, p. 1098).

Para tanto, apresenta-se o mapeamento dos processos da Figura 72.

Figura 72 - Mapeamento dos processos para construção do tesauro por Souza *et al*.



Fonte: elaborado pelo autor com base em Souza *et al*. (2014).

A idealização para construção do sistema envolveu, primeiramente, a criação de "[...] um grupo composto por profissionais do CPDOC e da Escola de Matemática Aplicada da FGV. Dentre as formações, historiadores, cientistas sociais, arquivistas, bibliotecários e cientistas da informação", isto é, "[...] caracterizando um ambiente de trabalho interdisciplinar, tal como a tarefa demanda" (SOUZA *et al*., 2014, p. 1101).

Enquanto a pesquisa de Vogel e Kobashi (2019) se aportou das metodologias orientadas pelo Arquivo Nacional da Austrália para construção de tesauro, Souza *et al*. (2014) serviram-se da norma ANSI/NISO Z.39.19-2005, desenvolvida pela American National Standard Institute/National Information Standards Organization

(ANSI/NISO, 2005). Os autores caracterizaram a metodologia preconizada pela norma (SOUZA *et al*., 2014, p. 1102) pelos seguintes passos:

1. Pesquisar os vocabulários existentes, de forma a evitar a duplicação e o retrabalho;
2. Determinar a estrutura e os formatos de exibição;
3. Escolher o método de construção, dentre as três possibilidades de abordagens: de comitê, empírica e indutiva;
4. Registro dos termos candidatos e seus relacionamentos (lógicos - hierárquicos e ontológicos - partitivos e associativos);
5. Validação dos termos;
6. Escolha dos termos candidatos (a fazer parte do produto);
7. Escolha dos níveis de especificidade;
8. Escolha de termos não associados (para garantir a estrutura do sistema).

De acordo com o mapeamento dos processos, visto anteriormente, foi delimitado o escopo do domínio investigado, bem como foram determinados os critérios de inclusão de termos no tesauro – para essa "[...] delimitação levou em consideração a organização atual das informações nos sistemas do CPDOC, o que ensejou a estratégia de elaboração de bases terminológicas integradas"; a ação da escolha do formato deu-se pelo tipo conceitual, com "[...] formas sistemáticas e alfabéticas"; já a escolha da abordagem para construção do tesauro, foi do tipo híbrida, com foco na abordagem comitê (SOUZA *et al*., 2014, p. 1102).

Em sequência aos processos, efetuou-se o registro dos termos e seus relacionamentos, os quais foram aprovados pelos analistas envolvidos; em seguida, para registro do tesauro,

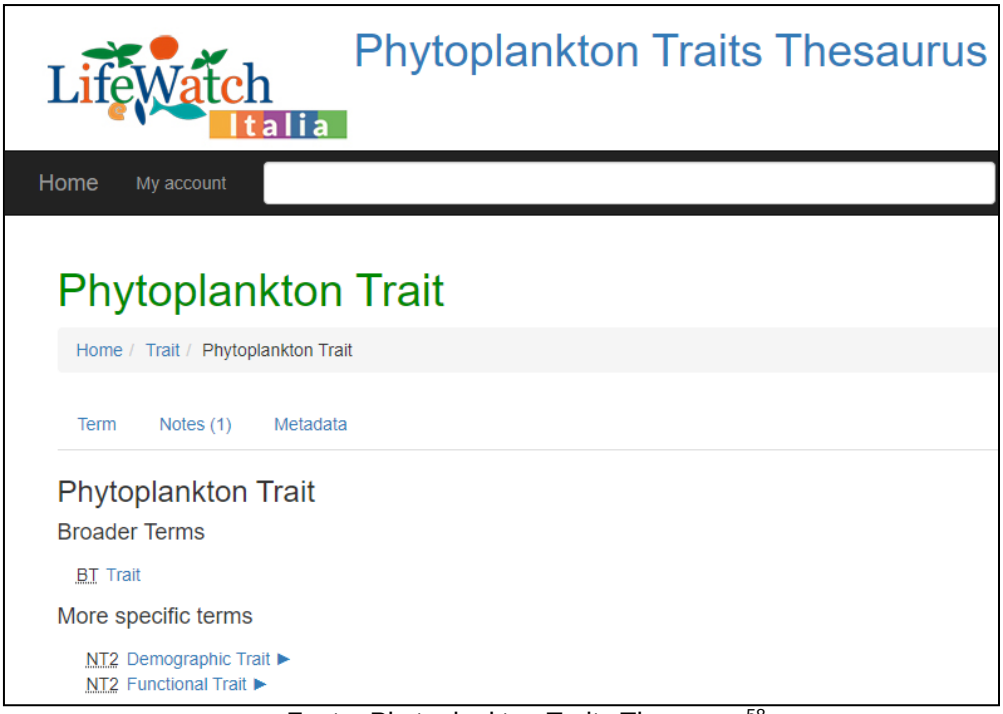
> [...] estudaram-se algumas ferramentas, como TCS10, que havia sido utilizado em trabalhos anteriores, mas descobriu-se que o produto foi descontinuado. Cogitou-se utilizar o software Protégé[55], utilizado para construção de ontologias, mas as ferramentas de suporte ao padrão SKOS[56], preconizado pelo W3C para publicação do tesauro e outros sistemas de organização do conhecimento no escopo da web semântica, ainda não se mostram satisfatórias. Após uma exaustiva pesquisa e teste de muitas das opções disponíveis, decidiu-se pelo uso do software Tematres[57] que, além de gratuito, é open source e oferece uma grande

[55] https://protege.stanford.edu/
[56] https://www.w3.org/2004/02/skos/
[57] https://r020.com.ar/tematres/demo/index.php

> gama de opções de exportação e visualização. (SOUZA *et al*. 2014, p. 1105)

O uso do software TemaTres é referenciado em outros tipos de tesauro, como, por exemplo, o tesauro PhytoTraits, um recurso semântico sobre as características funcionais do fitoplâncton – pesquisa de Rosati *et al*. (2017). O desenvolvimento do tesauro PhytoTraits, além de contar com o uso do TemaTres, também evidenciou a formação de uma equipe interdisciplinar. A seguir, uma representação do PhytoTraits.

Figura 73 - Tesauro Phytoplankton

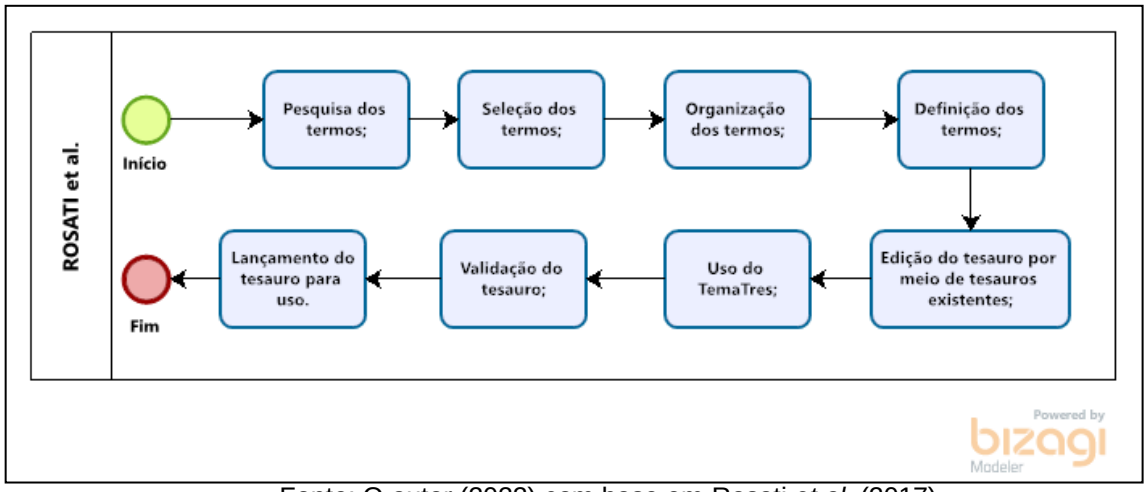


Fonte: Phytoplankton Traits Thesaurus[58]

Segundo os autores, esse tesauro "[...] é o resultado da colaboração interdisciplinar de especialistas do domínio funcional do fitoplâncton e das tecnologias de informação e comunicação[59]" (ROSATI *et al*., 2017, p. 129). Para construção do tesauro, a equipe seguiu as ações mapeadas a seguir e descritas na Figura 74.

[58] http://thesauri.lifewatchitaly.eu/PhytoTraits/index.php?tema=5&/phytoplankton-trait
[59] "PhytoTraits is the result of the interdisciplinary collaboration of experts both from the phytoplankton functional domain and from information and communication technologies" (ROSATI *et al*., 2017, p. 129).

Figura 74 - Mapeamento dos processos para construção do tesauro por Rosati *et al.*



Fonte: O autor (2022) com base em Rosati *et al*. (2017).

A primeira lista reuniu termos de um banco de dados e de um atlas de formas de fitoplâncton; posteriormente, outros termos relacionados foram extraídos de vocabulários, ontologias, pesquisas publicadas e outros documentos da equipe. Logo, “[...] todos os termos com significados semelhantes foram agrupados e um termo preferido foi escolhido”; por outro lado, “[...] sinônimos, grafias variantes e abreviações foram incluídos como termos não preferenciais”. Aqui, os autores argumentam sobre evitar redefinições, tomando como procedimento a organização dos termos “[...] em uma primeira estrutura hierárquica e fornecido com uma definição levando em consideração a literatura relevante e tesauros e ontologias pré-existentes”. (ROSATI *et al*., 2017, p. 134)[60].

As configurações do modelo SKOS deu base para o objetivo do reuso do sistema e “[...] para facilitar a edição do tesauro formatado em SKOS, o software de gerenciamento de tesauros de código aberto na web TemaTres foi escolhido” (ROSATI *et al*., 2017, p. 134). A justificativa do uso do TemaTres deve-se à simplicidade e à “[...] disponibilidade de sua interface web, ao mesmo tempo em que mantém um rico conjunto de recursos subjacentes”, o que “[...] tornou-o uma escolha adequada para o desenvolvimento de PhytoTraits” (ROSATI *et al*., 2017, p. 134)[61].

[60] All terms with similar meanings were grouped and a preferred term was chosen, while synonyms, variant spellings and abbreviations were included as non-preferred terms. Terms were organized into a first hierarchical structure and provided with a definition taking into account relevant literature and pre-existent thesauri and ontologies, to avoid redefinitions. (ROSATI *et al*., 2017, p. 134).

[61] In order to ease the editing of SKOS formatted thesaurus, the opensource and web-based thesaurus management software TemaTres (http://www.vocabularyserver.com/) was chosen (Gonzales-Aguilar *et al*., 2012). The relative simplicity of TemaTres and the availability of its web

Os autores relataram que “[...] durante a fase de edição, a lista inicial de termos, seus atributos e relações hierárquicas e associativas foram submetidos a diferentes modificações” (ROSATI *et al*., 2017, p. 135), como:

> [...] (i) modificou e enriqueceu os elementos estruturais de cada termo por meio da interface de usuário do TemaTres (termo preferencial, termo não preferencial, definição, nota bibliográfica associada, etc.), (ii) modificou a estrutura hierárquica e / ou (iii) adicionou ou excluiu um termo. (ROSATI *et al*., 2017, p. 135)[62]

Após as etapas descritas anteriormente, os autores submeteram o desenvolvimento do tesauro à validação de especialistas, “[...] que revisaram os termos e seus elementos estruturais, os aprovaram ou rejeitaram e propuseram modificações aos editores usando o elemento nota privada do TemaTres” (ROSATI *et al*., 2017, p. 135)[63]. Após as decisões dos autores quanto às observações feitas pelos especialistas, foi lançada a versão do tesauro para uso.

A atenção dada ao tratamento de termos é uma atividade peculiar, que exige da equipe de elaboração uma máxima de atenção. Tal como pode ser visto no tesauro conceitual de políticas de indexação para bibliotecas universitárias, publicado na pesquisa de Tartarotti, Dal’Evedove e Fujita (2015, p. 4), em que tratou da “[...] elaboração de um tesauro que exerça o papel de mediador entre os profissionais bibliotecários e usuários de bibliotecas universitárias, como forma de compartilhamento de termos”.

Essa máxima atenção dada ao tratamento de termos para elaboração de tesauros, por parte das autoras, pode ser vista no mapeamento dos processos apresentado em sequência.

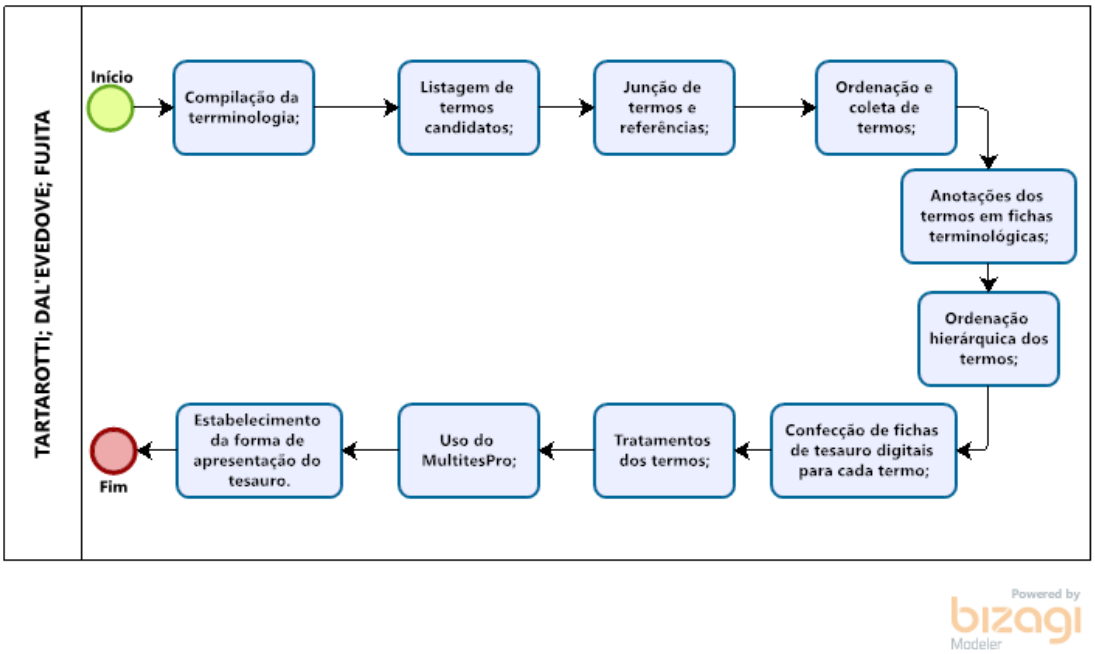
---

interface, while maintaining a rich set of underlying capabilities, made it a suitable choice for the development of PhytoTraits thesaurus.

[62] During the editing phase, the initial list of terms, their attributes and hierarchical and associative relationships were subjected to different modifications. Editors: (i) modified and enriched the structural elements of each term through the TemaTres user interface (preferred term, non-preferred term, definition, associated bibliographic note, etc.), (ii) modified the hierarchical structure, and/or (iii) added or deleted a term.

[63] Once the thesaurus was constructed, editors submitted it to the validators, who reviewed terms and their structural elements, approved or rejected them and proposed modifications to the editors by using the element private note of TemaTres. Editors decided whether to approve or not the validators' annotations and finally they released a new stable version of the thesaurus.

Figura 75 - Mapeamento dos processos para construção do tesauro por Tartarotti, Dal'Evedove e Fuijta



Fonte: O autor (2022) com base em Tartarotti, Dal'Evedove e Fujita (2015).

Orientado por uma máxima de tratamento de termos, conforme visualizado nos processos mapeados, o tesauro de Política de Indexação contou com três documentos bases, que constituíram o corpus de análise. A listagem dos termos candidatos foi feita por meio da "[...] leitura linha por linha de cada texto e digitação em arquivo do Excel de todos os termos, com exceção de palavras sem sentido". A etapa posterior baseou-se na "[...] frequência em que o termo apareceu nos textos e termos potencialmente sinônimos" (TARTAROTTI; DAL'EVEDOVE; FUJITA, 2015, p. 9-10). Assim, foi realizada a ordenação e coleta de termos com maior frequência.

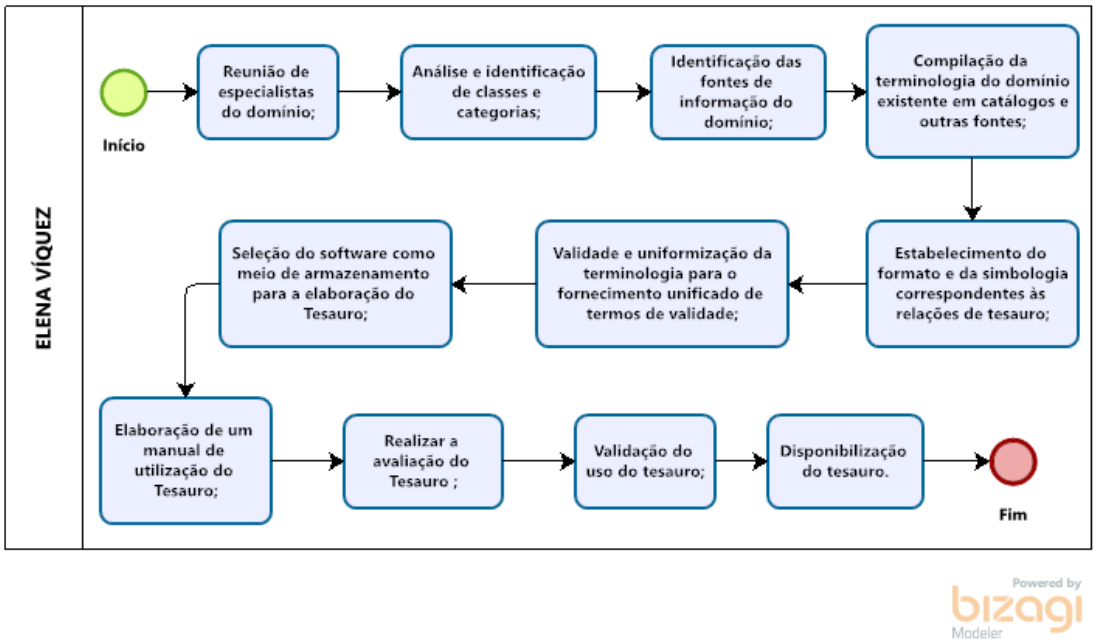
Concluída a etapa de ordenação e coleta, deu-se prosseguimento, por meio das "[...] anotações dos termos em fichas terminológicas de trabalho digitais", seguindo, assim, para a "[...] ordenação hierárquica dos termos em uma folha de distribuição, definindo-se a área, domínio e subdomínio" e a "[...] confecção de fichas de tesauro digitais para cada termo" (TARTAROTTI; DAL'EVEDOVE; FUJITA, 2015, p. 11).

A estruturação do sistema deu-se após a conclusão da etapa anterior. Para o gerenciamento terminológico, os autores utilizaram o software Multites Pro, com as percepções de que "[...] forneceu uma interface flexível e rápida e permitiu o

desenho da base terminológica em português, como também relatórios diferenciados". Logo, "[...] após os ajustes necessários, estabeleceu-se a forma de apresentação do tesauro, considerando-se o nível dos usuários, o que resultou na apresentação alfabética e hierárquica dos termos" TARTAROTTI; DAL'EVEDOVE; FUJITA, 2015, p. 11-12).

O uso de software para gestão de tesauro vem se tornando cada vez mais importante e acompanha, portanto, o aprimoramento desse sistema. *Multitest* um é dos exemplos de software para gerenciamento de tesauro. Conforme o tesauro mapeado anteriormente, Elena Víquez (2018, p.10) também relata o uso do *Multitest* para o Tesauro Especializado em Ciências do Movimento Humano. Contudo, a autora ressalta a necessidade de discutir, com a equipe responsável pelo desenvolvimento e pela manutenção do tesauro, qual o melhor software a ser adotado. Para o tesauro do Movimento Humano, foram discutidos quatro *softwares*: *ThManager*, *TemaTres*, *One-2-One* e o *Multitest*. Aqui, o *Multitest* recebeu destaque pela autora por relatar que "[...] foi recebido pela empresa *Multitest* treinamento para programar e fazer as correspondentes modificações no *site* em HTML", cujo armazenamento é no próprio tesauro. Apesar da avaliação dos *softwares* ter sido uma das principais etapas para desenvolvimento do SOC, essa etapa não iniciou o processo de construção do tesauro.

Figura 76 - Mapeamento dos processos para construção do tesauro por Elena Víquez



Fonte: O autor (2022) com base em Elena Víquez (2018).

Para iniciar o processo de desenvolvimento do tesauro, tal como Souza *et al*. (2014), Víquez (2018) reuniu uma equipe de especialistas do domínio para estruturar um plano do tesauro do Movimento Humano. A etapa de análise e identificação de classes e categorias foi feita por meio de uma revisão terminológica, identificando a equivalência de termos e os tipos de relacionamentos. Nesta fase, foi feito um estudo de abordagem mista, sendo quantitativo, por ser utilizado para medir, numericamente, categorias e termos, a quantificação do diagrama para construção do sistema, as fichas para controle terminológico, o número de fontes de informação, bem como a quantidade de especialistas envolvidos para revisão terminológica da área; e sendo, também, qualitativo, por analisar as diferentes relações constituídas dos termos que integraram as categorias do SOC. (VÍQUEZ, 2018).

As fontes de informação integraram especialistas em Ciências do Movimento Humano como estratégia para identificação de classes temáticas, Tal como categorias, vocabulários e notas de escopo e profissionais em Biblioteconomia, especialistas em desenvolvimento de tesauros. A participação ativa dos profissionais de Biblioteconomia também se deu para prova piloto do sistema. Foram citadas outras fontes de informação, como, por exemplo, documentários. (VÍQUEZ, 2018).

A análise mista também envolveu outras fases para análise qualitativa e quantitativa como instrumento de medição, por meio da utilização de

- Guia para identificação e validação das categorias que estruturaram o sistema;
- Tabela de identificação de fontes para coleta de dados;
- Formulário para controle terminológico;
- Diagrama de setas;
- Análise de ferramentas para elaboração de tesauros;
- Formulário de indexação para avaliação do sistema e validação para uso;
- Guia para elaboração de manuais para uso do SOC.

O uso de uma plataforma para armazenamento do sistema foi, conforme percebido, uma das ações que recebeu maior discussão. Quanto ao uso do *Multitest*, aponta-se que a "[...] plataforma tecnológica onde o tesauro é armazenado caracteriza-se por trabalhar com hipertexto nos descritores, o que permite a busca entre os mesmos descritores e vincula as buscas de descritores na web." (VÍQUEZ, 2018, p. 14).

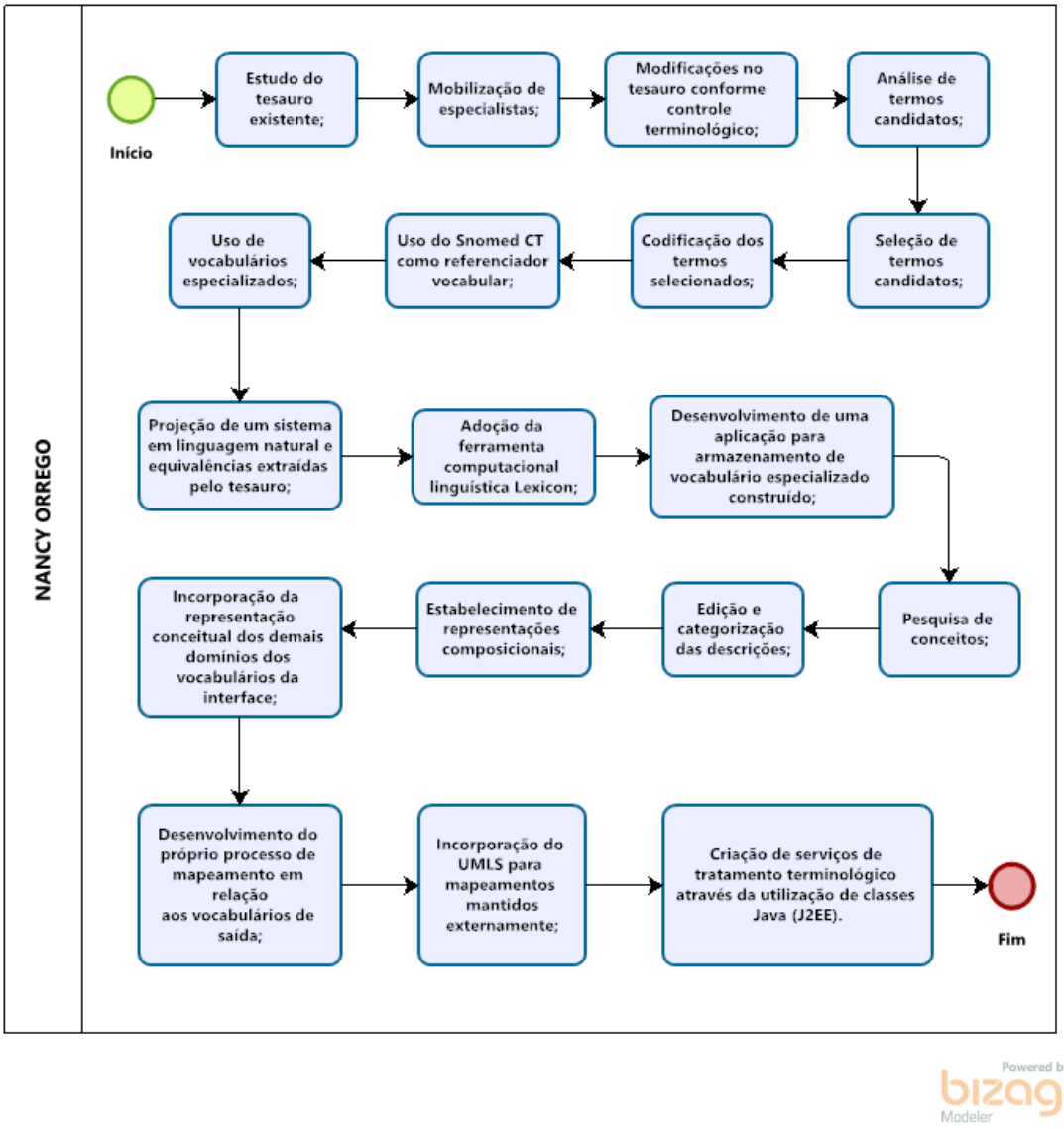
O emprego da tecnologia por profissionais de Ciência da Informação vem impactando fortemente os métodos e as técnicas utilizadas para estruturação de SOC, tendo o tesauro como um dos sistemas que recebeu maior atenção no que diz respeito ao aperfeiçoamento do sistema. Potentes em uso de tecnologia, o tesauro evoluiu em construção, funcionalidades e uso. Os exemplos dos tesauros vistos até aqui corroboram com as afirmativas. E, por fim, como forma de ilustrar de forma mais vertical, analisou-se dois outros tesauros: o primeiro, por representar o domínio da Saúde, especificamente aplicado em prontuários eletrônicos, uma pesquisa de Orrego (2018); outro, que trata sobre tesauro eletrônico, a importação no TemaTres e a disponibilidade desse sistema na *web*, por Santos, Cervantes e Fujita (2018).

Segundo a pesquisa publicada por Orrego, o projeto, no domínio da Saúde, surgiu

> [...] para resolver os problemas de interoperabilidade dos sistemas, sendo responsável pela criação, implementação e manutenção da estrutura central de vocabulários da instituição. Nos primeiros anos, ele usou diferentes técnicas para controlar o vocabulário, migrando de um sistema de listas de termos para a criação de um servidor de terminologia; resolvendo assim os aspectos semânticos, sintáticos e tecnológicos de interoperabilidade entre seus sistemas. (ORREGO, 2018, p. 92)

Para esse projeto, foram realizados os processos a seguir descritos na Figura 77.

Figura 77 - Mapeamento dos processos para construção do tesauro por Nancy Orrego



Fonte: elaborado pelo autor com base em Orrego (2018).

O destaque na aplicação de recursos tecnológicos para o mapeamento, refere-se, primeiramente, ao uso do Snomed CT[64], como referenciador vocabular.

> [...] escolhemos Snomed CT como vocabulário de referência. É uma nomenclatura desenvolvida pelo *American College of Pathologists*, há 30 anos evoluiu para diferentes versões, sendo a última "Snomed CT", por

[64] https://www.snomed.org/

> *Clinical Terms*, que resulta da união de "Snomed RT" com os "*Clinical Terms* V3" utilizado pelo sistema de saúde inglês. Esta nomenclatura, que abrange todos os campos da medicina, e tem uma estrutura organizada de conceitos e descrições, já está disponível para a língua espanhola e oferece uma funcionalidade de representação enumerativa e composicional, com definições semânticas dos conceitos por meio de interrelações entre eles. Snomed CT já foi mapeado com algumas classificações que podem ser usadas como vocabulário de saída e outros mapeamentos estão em construção no momento. (ORREGO, 2018, p. 83)

O Snomed CT busca desempenhar "[...] um papel essencial na melhoria da saúde da humanidade, determinando padrões para uma linguagem codificada que representa grupos de termos clínicos" e "[...] permite que as informações de saúde sejam trocadas globalmente para o benefício dos pacientes e outras partes interessadas". (SNOMED, 2021, *online*).

Em caráter do uso do Snomed CT por Orrego (2018, p. 84), ele fornece, ainda, "[...] uma lógica de composição robusta e explícita para a representação de vocabulários de interface. Desta forma, o vocabulário de referência, Snomed CT, atua como uma ponte entre o vocabulário do usuário e os vocabulários de saída".

> O vocabulário de referência deve ter um nível suficiente de detalhe e cobertura para representar os conceitos incluídos em todos os vocabulários de interface e vocabulários de saída. No caso do Snomed CT, as especificações para representação composicional permitem combinar termos para representar conceitos não originalmente incluídos na nomenclatura, atingindo percentuais de cobertura muito elevados. A estratégia que usamos para representar nosso vocabulário de interface no Snomed CT foi mapear os conceitos de cada um deles, um por um, para uma extensão Snomed CT criada por nós. Uma extensão é um conjunto de novos conceitos, descrições e relacionamentos, adicionados ao Snomed CT, cada conceito adicionado é descrito por seus relacionamentos com os conceitos padrão do Snomed CT. (ORREGO, 2018, p. 84)

Ainda, quanto à aplicação de recursos tecnológicos incorporados ao tesauro, cita-se a ferramenta de computação linguística Lexicon.

> Ao buscar os termos mais semelhantes ao texto que os profissionais inserem, a qualidade das opções encontradas pode ser melhorada com uma busca inteligente que considere as diferentes variações lexicais das palavras, como variações de gênero, número, inflexões, derivações, entre outras. Ele também pode conter as relações de sinonímia entre palavras diferentes. A ferramenta de computação lingüística que armazena essas informações e permite que sejam utilizadas em aplicações chama-se Lexicon. Um Lexicon contém tabelas com as diferentes variações das palavras conhecidas e regras que permitem que essas variações sejam inferidas a partir das características das palavras desconhecidas. Atualmente estamos desenvolvendo um Léxico que possibilite essas funcionalidades com as palavras utilizadas em nosso ambiente. A título de exemplo do funcionamento de um Léxico, ao inserir o texto "Pele irritada", a

> pesquisa obteria como opção “Irritação da pele”, que não seria encontrada numa simples pesquisa comparando o texto. (ORREGO, 2018, p. 87)

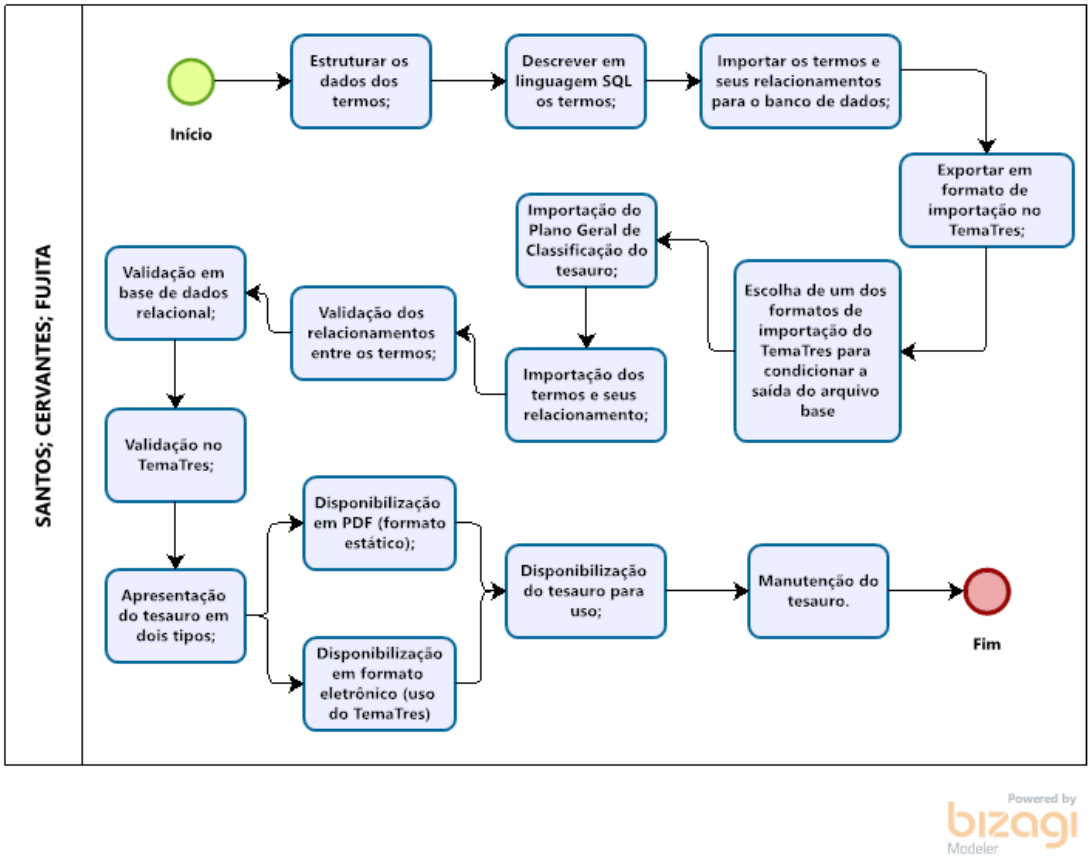
O uso dos recursos tecnológicos já existentes ainda não foi suficiente para agregar valor ao tesauro. Para tanto, foi desenvolvida uma aplicação, com o objetivo de manter o sistema no servidor terminológico. Neste caso, a preocupação se deu pela tarefa de manter o sistema e suas funcionalidades, o qual requereu à customização de um software passível de novas entradas, de edição dos conceitos e descrições, e de representação conceitual (ORREGO, 2018).

Mesmo utilizando o Snomed CT, a possibilidade de utilizar o Unified Medical Language System (UMLS) – Sistema Unificado de Linguagem Médica – não foi descartada. Como próxima etapa, Orrego (2018, p. 91) considerou a “[...] incorporação do UMLS para mapeamentos mantidos externamente”, bem como “[...] a criação de serviços de tratamento terminológico para as restantes aplicações de sistemas de informação, através da utilização de classes Java (J2EE)”.

Tesauros podem elaborados a partir de um planejamento bem definido, por meio dos membros da equipe de desenvolvimento do sistema, assim como podem ser revisados e são suscetíveis de manutenção, considerando a dinâmica da atualização científica, cultural e tecnológica. A ideia primária de construção de tesauros, fala-se aqui dos tesauros eletrônicos estáticos, recebe motivação de uso, hoje, por meio de técnicas e métodos de transformação desse sistema, integrando-o às plataformas de gerenciamento.

Tal como a pesquisa de Santos, Cervantes e Fujita (2018, p. 1110), que se basearam na “perspectiva de integração da sistemática de um vocabulário controlado, especificamente tesauro, em um sistema computacional gerenciador”, os pesquisadores tiveram o objetivo de “propor procedimentos metodológicos para conversão de tesauro eletrônico, formato estático, em tesauro eletrônico em um *software* gestor de tesauros e tornar disponível como recurso de busca e navegação dos termos do instrumento”. Esses procedimentos estão mapeados na Figura 78.

Figura 78 - Mapeamento dos processos para conversão de tesauro estático para um software gestor de tesauro por Santos, Cervantes e Fujita



Fonte: O autor (2022) com base em Santos, Cervantes e Fujita (2018).

A conversão do tesauro eletrônico estático, em PDF, para um *software* gestor de tesauro, deu-se por meio de uma análise entre o *software MultiTes* e *TemaTres*, dois softwares gestores já caracterizados nos mapeamentos vistos nesta seção. Para tanto, os pesquisadores optaram por converter o tesauro em PDF para o software gestor de tesauro *TemaTres*.

Segundo Santos, Cervantes e Fujita (2018, p. 1118), a primeira etapa, conforme visto na figura 76, foi estruturar os termos do tesauro explorado – o Tesauro Brasileiro de Ciência da Informação (TBCI). Esta etapa refere-se à estruturação dos termos do TBCI, verificando a modalidade, o formato e a estrutura dos termos, com a intenção de "[...] traçar um planejamento de como organizar os termos em tabelas, planilhas, arquivos textos ou outras formas que deem suporte eficaz e eficiente para segunda ação". A ação posterior deu-se pela descrição dos dados em linguagem SQL. Os autores apontam que a linguagem SQL "[...] é a forma padronizada de manipulação de dados no banco de dados" e que para esta

atividade utilizaram “[...] o comando INSERT para descrição dos termos dos vocabulários controlados e seus relacionamentos”.

> Para alcançar a descrição em linguagem SQL dos termos do TBCI, algumas atividades foram necessárias. Foi adicionada, para os termos, uma identificação. Nos campos USE, UP, TG, TR, TE e CAT, tem-se, em sua estrutura, um ou mais termos especificados para cada termo preferido e, por isso, foi necessário estabelecer um vínculo com o termo preferido por meio do seu identificador. Nos campos USE, UP, TG, TR e TE, que são termos que constam no rol de termos já identificados juntamente com os termos preferidos, foi necessário localizar cada um e adicionar o identificador para estabelecer os vínculos. Por exemplo, o termo AACR1, com seu identificador 1, tem em seu campo USE o termo Código de Catalogação Anglo-Americano cujo identificador é 380, portanto foi gerada uma coluna com o identificador do termo principal e outra coluna com o identificador do termo USE. Para este processo, foi desenvolvido um algoritmo para realizar o procedimento de maneira mais rápida e com risco zero de erros.
> Nesse algoritmo, foram utilizados os seguintes vetores: um vetor com os termos preferidos e não preferidos, outro com seus identificadores e um vetor com os termos. No exemplo utilizado do campo USE, o algoritmo percorre todo o vetor de termos preferidos e não preferidos até localizar o termo do campo USE. Ao localizar, gera-se o comando de insert utilizando a linguagem SQL, própria de manipulação de dados em sistemas gerenciadores de banco de dados, INSERT INTO TBR_USE (Id_termo, Id_Termo_use) VALUES (1, 380). Este processo se repete para os demais campos. (SANTOS; CERVANTES; FUJITA, 2018, p. 1120)

Ainda nesta fase, os autores contaram com o apoio da ferramenta Excel, da Microsoft. A descrição, na íntegra, do uso dessas técnicas, pode ser vista em Santos, Cervantes e Fujita (2018).

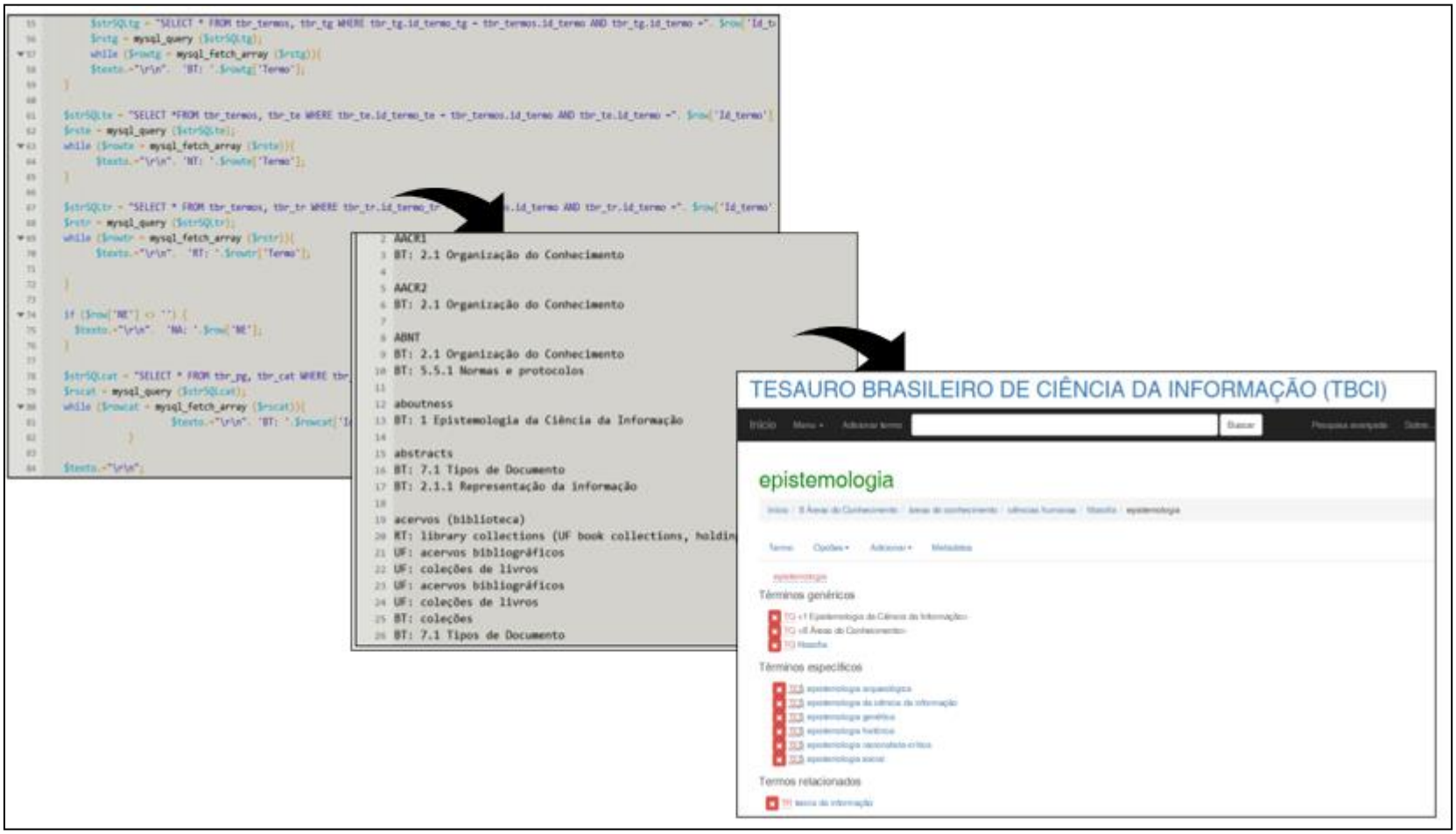
Prosseguindo, as ações da conversão do tesauro, na etapa de importação dos dados dos termos do vocabulário controlado, executaram-se as seguintes tarefas, conforme Santos, Cervantes e Fujita (2018, p. 1121):

> 1) Execução de todos os *SQLs*, para inserção do Plano Geral de Classificação na tabela denominada *TBR_PG*;
> 2) Execução de todos os *SQLs*, para inserção dos termos na tabela *TBR_TERMOS*;
> 3) Execução de todos os *SQLs*, para inserção dos vínculos dos termos principais com o campo USE na tabela *TBR_USE*;
> 4) Execução de todos os *SQLs*, para inserção dos vínculos dos termos principais com o campo UP na tabela *TBR_UP*;
> 5) Execução de todos os *SQLs*, para inserção dos vínculos dos termos principais com o campo TG na tabela *TBR_TG*;
> 6) Execução de todos os *SQLs*, para inserção dos vínculos dos termos principais com o campo TE na tabela *TBR_TE*;
> 7) Execução de todos os *SQLs*, para inserção dos vínculos dos termos principais com o campo TR na tabela *TBR_TR*;
> 8) Execução de todos os *SQLs*, para inserção dos vínculos dos termos principais com o campo CAT na tabela *TBR_CAT*.

Esse processo de importação objetivou a validação da estrutura do tesauro. Logo após esta etapa, foi possível gerar diversos formatos de arquivos de importação em SI. Para tanto, foi necessária a construção e o uso de algoritmos, para gerar de maneira automatizada; “[...] foi desenvolvido um algoritmo para gerar o arquivo no formato TXT, etiquetado com as estruturas do tesauro. O programa de computador foi desenvolvido utilizando a linguagem PHP para Web”; assim, “[...] a partir da execução faz-se o download do arquivo TXT gerado e a submissão ao processo de importação no TemaTres” (SANTOS; CERVANTES; FUJITA, 2018, p. 1122).

A figura 77 representa as fases: (1) codificação em PHP do algoritmo que gerou o arquivo etiquetado; (2) o arquivo etiquetado de importação; e (3) a estrutura sistemática do TBCI. Além de todos esses processos, ainda se somam outras duas ações: a primeira “[...] em uma base de dados relacional, como forma de validação dos relacionamentos dos termos” e a segunda no uso do TemaTres “[...] totalizando a importação de 2.058 termos, 1.828 relacionamentos entre os termos, 744 termos não preferidos e 336 notas de escopo” (SANTOS; CERVANTES; FUJITA, 2018, p. 1125).

Figura 79 - Codificação em PHP do algoritmo que gera o arquivo etiquetado de importação; e a estrutura sistemática do TBCI



Fonte: O autor (2022) com base em Santos, Cervantes e Fujita (2018).

Os autores listaram uma sequência de vantagens e debilidades sobre a relação entre as duas versões do tesauro do TBCI, o eletrônico estático em PDF e o eletrônico por meio do TemaTres, quanto aos processos de **busca**, **navegação** e **manutenção**. Santos, Cervantes e Fujita (2018, p. 1125-1126) explicam que

> A busca no formato estático é mais lenta, e o usuário deve saber exatamente a sintaxe do termo para utilizar o recurso dos leitores de arquivos PDFs, o comando Ctrl+F. Após a localização do termo, caso seja necessária a verificação de um dos termos com algum relacionamento deste primeiro termo localizado, será preciso realizar outra pesquisa. A navegação é caracterizada como sequencial para localização do termo a ser consultado, já que os termos se encontram organizados por ordem alfabética. Por outro lado, a manutenção do tesauro torna-se mais onerosa. Além da complexidade da estrutura relacional e terminológica que o instrumento proporciona, é necessário todo cuidado para realizar as atualizações de maneira que não quebre a sistemática do tesauro e consequentemente inconsistências nos relacionamentos dos temos.
> No formato disponibilizado no TemaTres, a busca de termos é realizada de três formas: pesquisa por termo com recurso de autocomplete; o usuário ao digitar o termo, a cada caractere incluído o sistema de busca localiza e atualiza a listagem dos termos, em que é possível escolher nesta listagem ou manter o termo digitado e executar a busca. A lista sistemática é a segunda forma de localizar um termo de forma hierárquica. A lista alfabética é a terceira organização para localizar um termo. Em decorrência da busca, a navegação é realizada por meio de links, em que cada termo é apresentado com sua url de acesso, o trânsito do usuário pelos termos é realizado de forma dinâmica e amigável. Na manutenção do tesauro pelo profissional da informação, é inexistente a preocupação com a quebra da sistemática do instrumento devido ao gerenciamento que o software proporciona. Diante desta possibilidade, a atividade do profissional que fará as manutenções ficará centralizada única e exclusivamente nos processos cognitivos e terminológicos.

Observa-se, de acordo com os mapeamentos acima, que a aplicação das Normas ANSI/NISO Z39.19 e ISO 25964 são regulamentações seguidas por pesquisadores e profissionais. A primeira, foi publicada pelo *National Information Stantdars Organization*, nos Estados Unidos, e buscou apresentar diretrizes para conteúdo, exibição, construção, testes, manutenção e gerenciamento de vocabulários controlados monolíngues; a segunda, ISO 25964, também conhecida como ISO 2788:1986, sediada em Genebra, apresenta os padrões para construção desse sistema e a interoperação com outros, fornecendo recomendações para o desenvolvimento e a manutenção, destinados a atividades de recuperação da informação. Ainda para esta Norma, existe uma segunda publicação, feita em 2013, que apresenta orientações para estabelecimento e manutenção de mapeamento entre tesauros ou entre tesauros e outros vocabulários (ASSIS; MOURA, 2016; BERTIR JUNIOR *et al*., 2017; CAMPOS; CAMPOS; SOUZA, 2017; GABRIEL

JUNIOR; LAIPELT, 2017; GABRIEL JUNIOR; LAIPELT, 2020; GONZALEZ AGUILAR, 2013).

Contudo, ao analisar as perspectivas abordadas pelas Normas acima, percebe-se o aperfeiçoamento de tesauros, em razão do aprimoramento das tecnologias digitais, conforme citado por Gabriel Junior e Laipelt (2017, p. 125-126), em busca de otimização para representação de tesauros

> [...] focados principalmente nos padrões da web semântica constituídos pelos padrões RDF e RDFs, surge em meados da década de 2010 o padrão Simple Knowledge Organization Systems (SKOS), como modelo para representação tanto para seres humanos como para computadores. Modelo este recomendado pela World Wide Web Consortium (W3C) para compartilhamento de dados vinculados. (...) Diferente das normas ISO e NISO, que não especificam de forma detalhada o estabelecimento das relações e possibilitam apenas o uso de relações simplificadas, ou seja, relações semânticas fracas, o SKOS possibilita o detalhamento dessas relações, e ainda a incorporação de novas relações por meio de novas classes de propriedades, com o uso do RDF.

Os softwares gerenciadores de tesauros foram, por assim dizer, destaques quanto ao incremento de tecnologias digitais ao desenvolvimento de tesauros. Por esta razão, a figura abaixo apresenta ferramentas – ou gerenciadores de tesauros, com base em Gabriel Junior e Laipelt (2017).

Na Figura 80, têm-se a nomenclatura dos gerenciadores e uma descrição de suas particularidades, que podem facilitar o processo de tomada de decisão por pesquisadores e profissionais quanto a melhor adoção do gerenciador em razão de tesauro.

Figura 80 - Ferramentas para construção de tesauros

Fonte: O autor (2022) com base em Gabriel Junior e Laipelt (2017, p. 132 – 134).

A elaboração do quadro sobre as concepções conceituais, acerca de tesauros, permite refletir que as condições conceituais, funcionais e operacionais passaram por aprimoramentos desde o momento do planejamento do sistema às suas condições de uso. Isso não renega outras concepções já existentes, porque, para se construir SOC, “[...] é necessário experimentar posturas teóricas mais condizentes com nosso tempo” (SALES, 2021, p. 15). Essa caraterização corrobora que, rapidamente, de acordo com o intervalo temporal estabelecido, o sistema tesauro evoluiu em todas as suas instâncias e pode ser apontado como um recurso potencial para atividades orientadas à organização e representação da informação e do conhecimento.

As concepções sobre tesauros é um fator importante para estabelecer que esse sistema pode ser lido e caracterizado de acordo com as suas perspectivas de construção e uso. Isto é, tesauro pode ser visto como instrumento, método, ferramenta, recurso, vocabulário e linguagens. Ainda pode-se acrescentar outras visões – neste caso, cabe ao elaborador, ou à equipe de elaboração, identificar os objetivos aos quais o sistema precisará atender.

Foram percebidos pontos de atenção quanto a tesauros e que, muitos deles, dependem fortemente do incremento de tecnologias digitais, para designar funcionalidades e operacionalidades. A aplicação e a resolução de pontos de atenção, nas funções e operações de tesauro, foram apresentadas nas duas instâncias a que esta tese se propôs: conceitual e processual.

Os processos mapeados são outros elementos que corroboram o aprimoramento de tesauros em razão do aperfeiçoamento tecnológico. A adoção de técnicas para extração textual ou terminológica, a aplicação de linguagens para compartilhamento de tesauros, os quesitos voltados à interoperabilidade, o estudo e determinação de gerenciadores de tesauros, são alguns dos constituintes e referenciais desse aprimoramento.

Contudo, apesar das reflexões quanto aos pontos de atenção que carecem de soluções tecnológicas, outros têm sido demandados e em caráter de urgência.

A ações de organização e representação da informação e do conhecimento requerem conhecimento de princípios, teorias e diferentes padrões e práticas nacionais e internacionais; aqui, também se evidenciam as necessidades de criar e compartilhar informação e que, para isso, necessita-se de conhecimento de

ferramentas, estruturas para armazenamento, compartilhamento, e acesso às informações. A educação de pesquisadores e profissionais da área de organização e representação da informação e do conhecimento não se pode limitar-se à "[...] catalogação ou aos padrões de metadados gerais e específicos do domínio. Para formar um profissional bem-informado, é necessário cobrir o contexto mais amplo" (SALABA, 2020, p. 384). Esse contexto pode ser, portanto, caracterizado por classificação, indexação, modelos conceituais, divulgações de pesquisas, contribuindo para estudo de melhores práticas, processamento e gerenciamento de dados etc. No sentido de harmonizar as percepções teóricas e práticas acerca do sistema tesauro, foi necessário equilibrar a tesa entre as discussões conceituais naquele primeiro momento e, em seguida quanto a aplicação daqueles conceitos nos moldes processuais.

Um outro ponto de atenção para estudos e planejamentos de tesauros, somado ainda ao aprendizado de pesquisadores e profissionais de organização e representação da informação e do conhecimento, é o aprendizado relacionado às considerações éticas, consciência intercultural, a capacidade de promover sistemas que atendam às necessidades do domínio, bem como ao de seus devidos usuários, às tendências globais, e à colaboração e ao compartilhamento de informação e conhecimento (SALABA, 2020). Pinho e Milani (2021, p. 204), ao tratarem desses pontos, apontam que profissionais e pesquisadores "[...] devem ter consciência das barreiras que impedem uma representação bem próxima da crença do usuário". Tão logo, ainda conforme os autores, "[...] o futuro dos tesauros, além de estarem atrelados ao fator de ajuda interdisciplinar, ainda estão nas mãos de seus idealizadores". A isso se deve o fato de que "[...] as questões que envolvem os aspectos culturais e os domínios científicos são discussões que permanecem a cargo desses profissionais", aos quais devem se orientar às conformidades éticas e consciências interculturais para disseminação e acesso à informação.

4.3 Taxonomia: conceitos e processos

Como visto na seção 3 – Sistemas de Organização do Conhecimento, o advento conceitual do termo "taxonomia" pertence ao domínio da Biologia, tendo sido utilizado pela primeira vez por Carlos Lineu, no século XVIII. Posteriormente, a

Pedagogia apropriou-se dele, quando Benjamim Bloom o aplicou para classificar os objetivos educacionais. A Informática, então, adotou-o para uso em meio eletrônico e, assim, levou o conceito do sistema para a Ciência da Informação – mais especificamente para os estudos de organização e representação da informação e do conhecimento.

A taxonomia, na Ciência da Informação, é vista como um dos Sistemas de Organização do Conhecimento rico em aplicação em atividades plurais, partindo da indexação de assuntos ao acesso e recuperação da informação. Com o passar do tempo, o conceito desse sistema foi aperfeiçoado, trazendo consigo novas práticas e técnicas para seu desenvolvimento e aplicação.

De acordo com o quadro bibliográfico explorado, foram mapeadas as concepções conceituais e processuais apresentadas pelos autores devidamente referenciados. O Quadro 10 abaixo aborda o sistema taxonomia.

Quadro 10 – Concepções conceituais sobre taxonomia.

ANO	AUTOR(A)(ES)(AS)	CONCEPÇÕES CONCEITUAIS
2010	AGANETTE; SOUZA; ALVARENGA	As taxonomias apresentam definições variadas, assim como suas características, origens e usos práticos em diferentes ambientes informacionais. Entretanto, algumas dessas características, usos e aplicações são reincidentes, isto é, são colocadas por diferentes autores. Uma delas é a construção sistemática e a ordenação da informação a partir de sua estrutura hierárquica de subordinação de assuntos. Outro ponto em comum, e bastante ressaltado, diz respeito à taxonomia ser um tipo de vocabulário controlado.
2010	NOVO	O conceito de taxonomia pode ser definido, atualmente, como estruturas classificatórias para organização de domínios de conhecimento. Tem por finalidade permitir agregação de informação e dados, possibilitando acesso através de navegação. No âmbito dos estudos da Ciência da Informação, as taxonomias estão sendo vistas como ferramentas de RI, além de enfatizar o aspecto navegacional. Preferimos evidenciar seu aspecto de ferramenta classificatória de um dado domínio.
2011	FITTERER; METLER RHONER; WINTER.	Taxonomia se mostra útil para integrar considerações estratégicas com capacidades técnicas/funcionais baseadas nos requisitos das partes interessadas expressos pelos critérios da taxonomia.[65]

[65] taxonomy thereby proved useful to integrate strategic considerations with technical/functional capabilities based on the stakeholders' requirements expressed by the criteria of the taxonomy

2012	PONTES; LIMA	Taxonomias são estruturas hierárquicas de categorias, desenvolvidas de forma sistemática, utilizadas como ferramenta de organização intelectual. São, geralmente, empregadas em portais institucionais e bibliotecas digitais, como um mecanismo de consulta, ao lado de ferramentas de busca [...]. A organização das informações, através de uma taxonomia, permite alocar, recuperar e comunicar informações dentro de um sistema de maneira lógica, pela navegação. As taxonomias são, geralmente, desenvolvidas utilizando a abordagem enumerativa, através da segmentação do universo, na qual as categorias são dispostas em uma estrutura hierárquica e os níveis das subcategorias são definidos. [...] as taxonomias facetadas, são projetadas de acordo com o modelo analítico-sintético [...] cada faceta consiste de uma terminologia, ou seja, um conjunto finito de termos estruturados por relação de subsunção (hierárquica). Cada faceta é projetada separadamente e modela um aspecto distinto do domínio. Com o uso de uma taxonomia facetada, cada objeto pode ser classificado a partir da seleção de um ou mais termos a partir de diferentes facetas. [...] O modelo para taxonomias dinâmicas somente requer uma taxonomia multidimensional (que permita a classificação de um objeto por mais de um conceito), que não precisa ser necessariamente facetada. O projeto baseado na classificação facetada contribui para a construção de taxonomias dinâmicas compactas e de fácil entendimento e uso pelos usuários.

2012	VITAL	A gestão da informação e do conhecimento encontra nos portais corporativos importante ferramenta para atingir seus objetivos, entretanto, o tratamento das informações e a codificação dos conhecimentos requerem o desenvolvimento e a aplicação de modelos de representação comuns a todos os usuários do sistema. A taxonomia, um modelo de representação de informação e conhecimento, vem sendo considerada uma metodologia facilitadora tanto da organização quanto da recuperação, representação e disponibilização do conteúdo em ambiente Web. [...] A taxonomia organiza a informação da mais genérica à mais específica, utilizando-se da relação gênero-espécie entre os termos. [...] A taxonomia descritiva consiste na criação de um vocabulário controlado com o objetivo de otimizar a busca e recuperação das informações. [...] A taxonomia de navegação é baseada nos modelos mentais de organização das informações. A estruturação das informações fica visível ao usuário final [...] A taxonomia para gerenciamento de dados é composta de uma pequena lista de termos autorizados sem qualquer estrutura hierárquica, assemelha-se com as taxonomias descritivas; não tem, todavia, a pretensão de representar e promover acesso a toda informação produzida na organização. [...] Pode-se ainda acrescentar outros três tipos de taxonomia: por assunto; por unidade; e taxonomia funcional.
2012	ALEXANDER	Taxonomia tem como sistemas relacionados: “esquemas de classificação formal, ontologias, vocabulários controlados, tesauros e linguagens de indexação [...] Taxonomias e dicionários são representações estruturadas da linguagem e do uso [...] A estrutura e a linguagem da taxonomia determinam, portanto, o contexto e, portanto, o significado da colocação dos itens. Portanto, a linguagem em taxonomias não reflete meramente preferências ou escolhas de estilo, mas afeta quem é capaz de entender e usar a taxonomia e quem é excluído" [...] às vezes, as taxonomias são consideradas instrumentos transparentes, lógicos e abertos de organização do conhecimento e isso pode até ser um objetivo expresso por seus criadores.[66]

[66] Related systems such as formal classification schemes, ontologies, controlled vocabularies, thesauri, and indexing languages [...] Taxonomies and dictionaries are both structured representations of language and usage. [...] The structure and language of the taxonomy thus determines the context and hence the meaningfulness of the placing of the items. Therefore language in taxonomies does not merely reflect preferences or style choices, but affects who is able to understand and use the taxonomy and who is excluded. [...] Taxonomies are sometimes assumed to be transparent, logical, and open instruments of knowledge organisation and this may even be an expressed aim of their creators

2014	ALEXANDER	As taxonomias são construídas por pessoas ou grupos específicos dentro de contextos organizacionais, sociais, históricos ou políticos particulares, por isso é provável que o ponto de vista dominante na época terá a influência mais forte sobre a estrutura e a terminologia usadas na taxonomia...[67]
2014	PIETERSE; KOURIE	Uma taxonomia é uma coleção organizada hierarquicamente que contém itens e seus atributos. [...] Uma taxonomia é criada agrupando coisas em um domínio em categorias e subcategorias. Frequentemente, as subcategorias são formadas em vários níveis de profundidade. Quando dois conceitos estão em uma relação hierárquica, o superconceito é denominado hiperínimo do subconceito, e o subconceito é denominado hipônimo do superconceito. Ele garante a transferência de propriedades de superconceitos para subconceitos.[68]
2014	SILVA; FREITAS; SILVA	O processamento semântico automático, por meio de taxonomias, tem se tornado cada vez mais comum nos últimos anos, pois fornece os meios para representar e usar o conhecimento do mundo. Em aplicações que envolvem o processamento de conteúdo textual, esse conhecimento de mundo pode significar, por exemplo, o entendimento sobre o conteúdo de um texto. Taxonomias podem ser usadas para identificar, em um texto original, seus principais tópicos para subsidiar as mais diversas áreas. [...]; taxonomia como uma das estratégias para identificar tópicos e como suporte para o desenvolvimento de software que utilize esse conhecimento.
2015	STEIMER; LUZ	Arquitetura da Informação e Experiência do Usuário são responsáveis pela identidade digital do website de comércio eletrônico. [...] Nessa estrutura informacional, a organização da informação se dá por meio da criação e do desenvolvimento das taxonomias. [...] Deste modo, em ambiente web, as taxonomias facilitam o acesso e navegação, contribuem para findability e para recuperação inteligente.

[67] Taxonomies are constructed by specific people or groups within particular organisational, social, historical, or political contexts, so it is likely that the dominant viewpoint at the time will have the strongest influence on both the structure and the terminology used in the taxonomy...

[68] A taxonomy is a hierarchically organized collection that contains items and their attributes. [...] taxonomy is created by grouping things in a domain into categories and sub-categories. Often sub-categories are formed several levels deep. When two concepts are in a hierarchical relationship, the super-concept is called the hypernym of the sub-concept, and the sub-concept is called the hyponym of the super-concept. It ensures the transfer of properties from super-concepts to sub concepts.

2015	WU, YANG	A taxonomia foi construída para apoiar a descoberta de conhecimento por meio de inferência, não para organizar e navegar em recursos de informação. [...] O desenvolvimento de padrões de inferência específicos e sistemáticos para descoberta de conhecimento pode ser um tópico para estudo futuro.
2016	SIMÕES; FREITAS; GRACIOSO; BRAVO	Com efeito, em diversas definições de taxonomia encontra-se explícita a ideia de que se trata de uma classificação. [...] Embora seja considerada uma classificação, a taxonomia apresenta características próprias e distintivas, tais como: o seu uso em ambientes digitais e em contextos de Web Semântica; o seu uso alto nível de organização; a sua estruturação usual em linguagem vocabular e, consequentemente, não notacional; a sua flexibilidade e o seu dinamismo. [...] No âmbito da ciência da informação, verifica-se em geral a presença de esquemas taxonômicos associados à organização e à recuperação da informação e, por extensão, ao conhecimento; em particular às linguagens documentais. Inicialmente, a sua presença faz-se sentir nos ambientes tradicionais, entre os quais as bibliotecas convencionais. Posteriormente, e acompanhando a evolução do próprio campo, seus princípios e métodos estendem-se aos, então recém-nascidos, ambientes virtuais, nos quais se inserem os portais e as bibliotecas digitais. [...], desde o final do século XX, concretamente até o advento das novas tecnologias, este tipo de linguagem foi uma das mais utilizadas, quer na organização do conhecimento, quer na permuta de informação.
2016	SANTOS; MARTINS; KANEOYA	Taxonomia como importante instrumento de representação na Inteligência competitiva; diversos são os projetos de inteligência competitiva que podem utilizar taxonomias como forma de apoiar o desenvolvimento das atividades
2017	CORRÊA; LACERDA; ZIVIANI; FRANÇA; RIBEIRO	[...] Taxonomias consistem na categorização de diversos conceitos recortados por similaridade aos quais eles servem.

2017	SOUSA; ARAÚJO JÚNIOR	Tradicionalmente, as taxonomias tiveram por função a classificação das espécies em botânica e zoologia, adotando uma nomenclatura binária. O resgate da taxonomia nos sistemas de informação considera a unidade sistemática (taxon) não mais família, gênero ou espécie, mas conceitos. [...] A taxonomia pode ser considerada uma estrutura que possibilita classificar objetos, seres vivos, coleções de livros ou documentos, em grupos ordenados hierarquicamente, a fim de possibilitar sua identificação, localização e acesso. Pode também ser definida como um sistema de classificação que apoia o acesso à informação, permitindo alocar, recuperar e comunicar informações em um sistema, de maneira lógica.
2017	CUNHA; ARAÚJO JÚNIOR	A taxonomia pode ser conceituada como uma linguagem documentária, que possui um modelo de classificação voltado para a operacionalização do processo de busca e recuperação da informação, e tem como como objetivos: organizar e representar conceitos por meio de termos; padronizar a comunicação organizacional; e buscar consenso na comunicação e classificação dos temas da organização, por meio da padronização dos termos usados para descrever e classificar os assuntos. [...] As taxonomias podem se constituir em ferramentas efetivas na padronização do registro do conhecimento, por meio do uso de uma linguagem comum (rótulos) ao descrever métodos e processos organizacionais e a organização das informações das bases de conhecimento em portais corporativos. O registro do conhecimento deve trazer uma série de benefícios na conversão de conhecimento tácito em explícito e entre estes benefícios podemos destacar aqueles que poderão se valer das potencialidades das taxonomias na descrição padronizada de métodos e processos organizacionais.
2017	PLAZA-CARVAJAL; DUTRA; MACEDO	A elaboração de uma taxonomia [como ferramenta], em determinado campo do conhecimento, implica em uma visão sistêmica em relação à organização e à classificação dos seus componentes. Isso quer dizer que deve existir um mapeamento completo, sendo evidenciados os vínculos entre as diversas categorias que o compõem

2018	LOPES; AGANETTE; MACULAN;	As taxonomias são consideradas formas de organização e representação do conhecimento que buscam organizar a produção intelectual humana. O uso da taxonomia tem sido adotado por facilitar o acesso, pois a navegação se dá por meio de classe, subclasses e assim sucessivamente, intensificando o nível da especificidade. [...] Recentemente, o uso e desenvolvimento de taxonomias tem sido um dos pontos principais da gestão da informação e do conhecimento [...] Taxonomia corporativa tem uma estrutura que permite que as organizações identifiquem seus processos e atividades, na recuperação e acesso às informações diárias, diminuindo assim a duplicação de informações e melhorando os esforços, auxiliando assim a tomada de decisão. Por outro lado, a taxonomia facetada tem uma estrutura que facilita o entendimento de um conceito e o torna compreensível ao usuário. Para facilitar a recuperação dos dados, é necessária a construção de um vocabulário controlado. Esse vocabulário é dividido em diversas subclasses, formando, assim, uma lista de conteúdos de diferentes tipos que retratam diferentes aspectos de criação, esses aspectos são conhecidos como facetas, e esse modelo de vocabulário controlado é denominado de taxonomia facetada.

2018	LUZ	No mundo da informação digital a interface é o meio pelo qual o usuário se relaciona com as informações. Nesse contexto, a classificação tem uma grande importância, considerando a grande variedade de objetos informacionais disponíveis. A interface é um elemento básico do processo de organização e de acesso à informação; [assim] os metadados organizam as informações que serão acessadas por meio das interfaces digitais. Por sua vez, o conjunto de metadados disponibilizados em interfaces chama-se Taxonomia; e, para planejar a taxonomia digital, utiliza-se a Arquitetura de Informação, que, projetada de forma eficiente, agiliza a conclusão de tarefas executadas pelos usuários pelo conteúdo, levando em conta a navegação do usuário. Não se pode projetar a experiência do usuário, mas é possível visualizar como ela pode acontecer. Por exemplo, se o consumidor não encontrar o livro que procura, ele não pode obter a publicação. Geralmente quando usuários não encontram algo que querem em um site, eles acreditam que esta biblioteca não tem a publicação desejada. Portanto esta não pode ser uma boa experiência. Torna-se, portanto, um desafio melhorar a probabilidade de encontro por meio da padronização de recursos que possibilitam a experiência do usuário e por meio das técnicas de taxonomia.
2018	BRAGA	As taxonomias são aplicadas para portais institucionais, repositórios institucionais, Web Semântica, ontologias, gestão de informação (IM) e gestão de conhecimento (KM) como um novo motor de consulta ao lado das ferramentas de pesquisa tradicionais.; O conhecimento gerado pela taxonomia pode ser usado para facilitar processos de organização e recuperação de informações, bem como sua própria compreensão da coleção textual organizada, ou mesmo servir como suporte para sistemas de suporte à decisão. Mas é importante que as técnicas sejam desenvolvidas para auxiliar especialistas no domínio, a fim de facilitar a compreensão e o uso dos conhecimentos adquiridos
2019	CHENG; LUDASCHER	É uma estrutura hierárquica de árvore de termos (ou nomes) com cada nó tendo apenas um pai (com exceção do nó raiz não tendo pai). Cada nó é desconectado de seu nó irmão, o que significa que dois nós no mesmo nível na árvore são mutuamente exclusivos um do outro (regra de desconexão de irmãos). Os filhos em cada ramo da árvore também são considerados os únicos filhos do pai sem filhos pródigos...[69]

[69] is a hierarchical, tree structure of terms (or names) with each node having only one parent (with the exception of the root node having no parent). Each node is disjointing from its sibling node, meaning

2019	SALES; SAYÃO	Taxonomia possa ser útil na modelagem mais precisa de plataformas dedicados à curadoria de dados pesquisas e seus múltiplos propósitos; [taxonomia auxilia] no estudo, mapeamento e caracterização dos dados de pesquisa proporcionando uma modelagem de sistema de dados mais eficaz e uma definição precisa dos dados no âmbito da construção de políticas de gestão de dados.

Fonte: O autor (2022) baseado em dados da pesquisa.

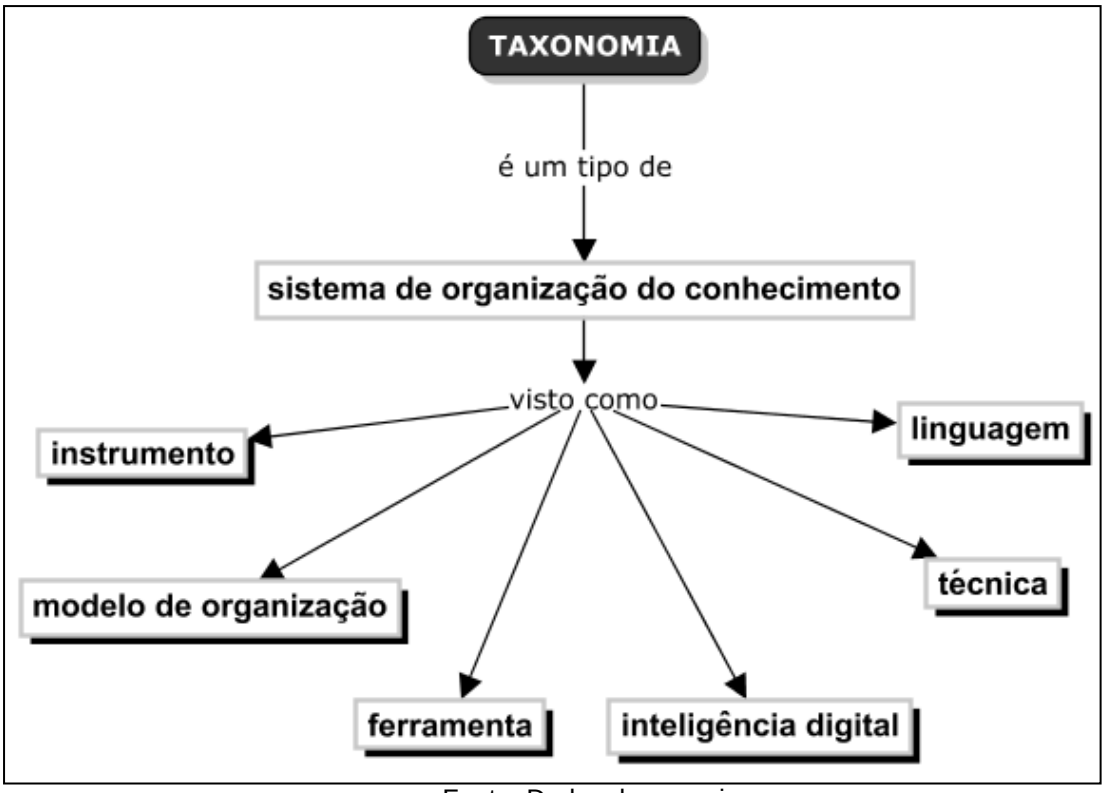
O Quadro 10 acima explora concepções conceituais sobre o sistema taxonomia. Nesse sentido, a reunião dos conceitos sobre taxonomia esclarece a riqueza conceitual em termos de definição, desenvolvimento, aplicação e uso; além de apontar contribuições tecnológicas digitais como fator determinante para a dinâmica operacional desse sistema (BRAGA, 2016; BRAVO, 2016; LUZ, 2018; NOVO, 2010; SILVA; FREITAS; SILVA, 2014; PONTES; LIMA, 2012; SIMÕES *et al*. 2016; STEIMER; LUZ, 2015; VITAL, 2012).

É fundamental ressaltar que não se pretende limitar as discussões conceituais sobre taxonomia, mas clarear as concepções sobre esse sistema na Ciência da Informação. Assim, o mapa conceitual apresentado a seguir reúne os extratos identificados no quadro acima.

---

that two nodes on the same level in the tree are mutually exclusive of each other (sibling disjointness rule). The children in each branch of the tree are also considered to be the only children of the parent with no prodigal sons...

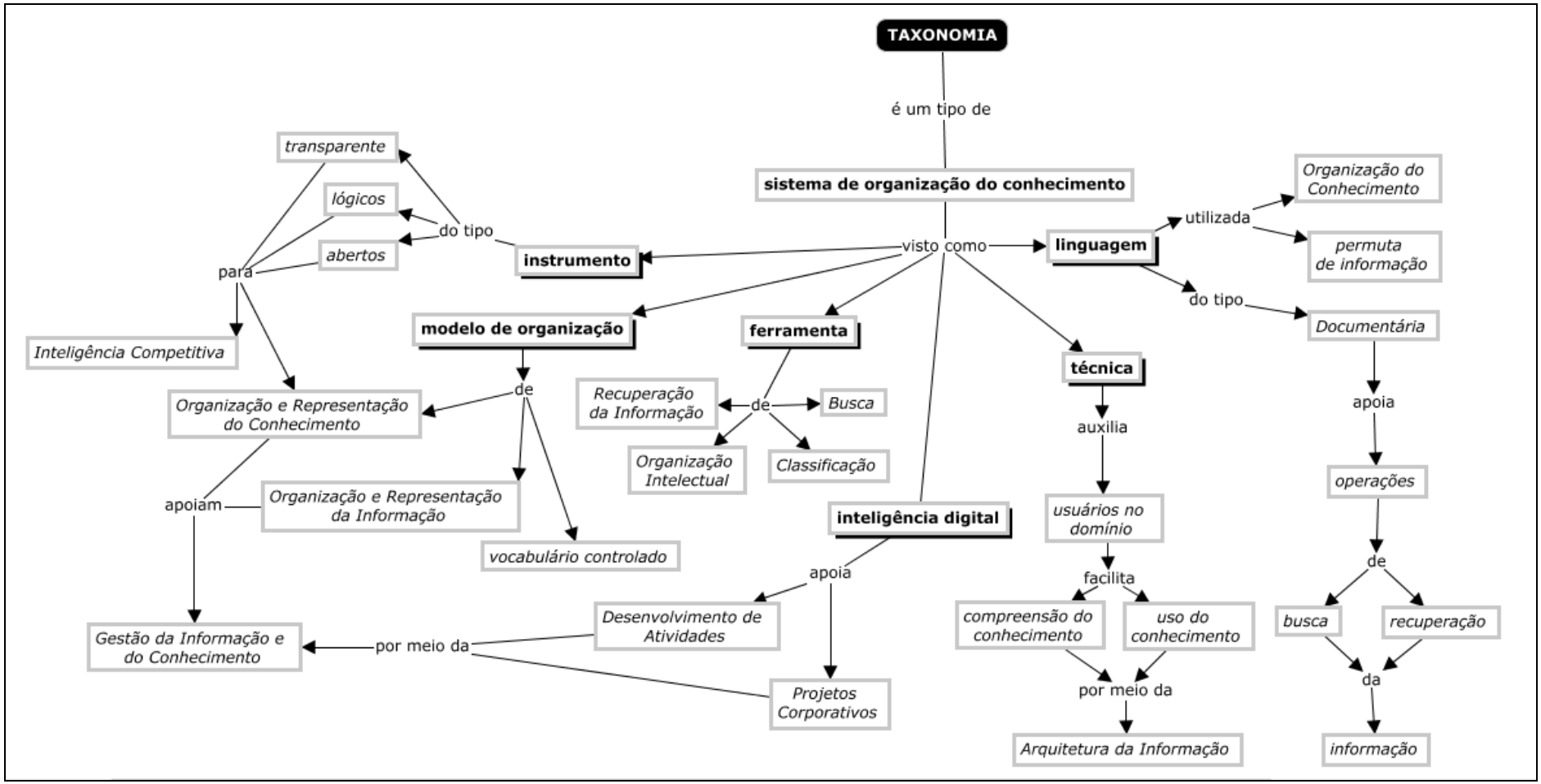
Figura 81 - Concepções sobre taxonomia



Fonte: Dados da pesquisa.

A leitura da ilustração acima implica compreender que taxonomia é um tipo de sistema de organização do conhecimento visto como: **instrumento** de apoio à organização e representação do conhecimento; **modelo de organização** para práticas de gestão da informação e do conhecimento; **ferramenta** de recuperação da informação; de classificação; de organização intelectual; de busca; para padronização de registros de conhecimentos; de pesquisas tradicionais; **inteligência digital** competitiva; **técnicas** de organização e representação da informação e do conhecimento; e **linguagem**. A extensão dessas concepções conceituais sobre taxonomia está apresentada abaixo.

Figura 82 - Concepções estendidas sobre taxonomia



Fonte: O autor (2022) baseado em dados da pesquisa.

Aganette, Alvarenga e Souza (2010, p. 77), ao realizarem um estudo sobre textos que tratavam de taxonomia, perceberam incidências conceituais as quais definiram o sistema como uma ferramenta, ou um processo, com diferentes origens, usos e funções – para a Ciência da Informação, pode ser visto como um sistema de classificação. Assim, o mapa conceitual elaborado acima buscou representar as ideias de um novo panorama que os estudiosos de Ciência da Informação têm, hoje, sobre o universo taxonômico. Isso se deve ao fato da multidimensionalidade operacional recebido por este sistema nos últimos anos, principalmente pelo aperfeiçoamento sistêmico em razão do aprimoramento das tecnologias digitais.

O mapa elaborado corrobora com a ideia de que a taxonomia pode ser vista como um tipo de instrumento lógico, o qual auxilia o planejamento organizacional orientado à Inteligência Competitiva (SANTOS; MARTINS; KANEOYA, 2016; STEIMER; LUZ, 2015), atraindo usuários à visitação de seus ambientes digitais, bem como orientando o negócio para uma melhor experiência do usuário. Podem-se citar, como exemplo de ambiente digital, os portais corporativos que, por sua vez, evoluíram no sentido de apresentar conteúdo específico para usuários definidos, facilitando a comunicação entre colaboradores e clientes (VITAL, 2012, p. 30).

Na reunião dos conceitos até aqui discutidos, pode-se inferir, ainda, que esse sistema é usado como uma técnica para auxiliar usuários: às operações de busca e recuperação da informação e à compreensão e utilização do conhecimento registrado por meio do desenho arquitetônico para tais processos.

Assim, no sentido de elucidar as concepções conceituais sobre este sistema, bem como relacioná-los em um escopo sobre tecnologia digital; gestão da informação e do conhecimento; e organização e representação da informação e do conhecimento, elaborou-se o esquema a seguir.

Figura 83 - Interseção entre Taxonomia e Organização, Gestão e Tecnologia

Fonte: O autor (2022) baseado em dados da pesquisa.

Explorar concepções conceituais sobre este sistema implicou considerá-lo como um dos sistemas de organização do conhecimento aperfeiçoado ao longo do tempo; isso porque, de acordo com os textos estudados, pode-se observar a complexidade da dinâmica operacional, ocasionada pelo uso de taxonomias em múltiplos ambientes, sobretudo nos digitais.

O aprimoramento tecnológico permitiu que o desenvolvimento e uso de taxonomia se tornassem mais facetados e dinâmicos. Pontes e Lima (2012), por exemplo, comentaram este sistema sob três prismas, a saber: a) taxonomias tradicionais, que utilizam uma abordagem enumerativa, adotando uma estrutura hierárquica de fácil compreensão e uso, porém não flexível; b) taxonomia facetada, a qual corresponde a uma estrutura menos rígida e mais facetada, onde cada objeto pode ser classificado por meio de seleção de um ou múltiplos termos, a partir de diferentes facetas; e c) taxonomia dinâmica, condizente à operacionalização deste sistema, considerando-o como um sistema de busca facetada. Para tanto, em um exemplo prático sobre esse tipo de taxonomia, as autoras explicam que

> O modelo para taxonomias dinâmicas somente requer uma taxonomia multidimensional (que permita a classificação de um objeto por mais de um conceito), que não precisa ser necessariamente facetada. O projeto baseado na classificação facetada contribui para a construção de taxonomias dinâmicas compactas e de fácil entendimento e uso pelos usuários. Outra vantagem observada diz respeito à estabilidade e, como consequência, a pouca necessidade de manutenção da taxonomia, além de proporcionar uma estrutura mais enxuta, intuitiva e com melhor capacidade de convergência. (PONTES; LIMA, 2012 p. 31)

A concepção conceitual de Vital (2012), vista no Quadro 10, sobre possíveis classificações do sistema taxonômico, como: descritiva; de navegação; e para gerenciamento de dados (CONWAY; SLIGAR, 2002), trata de outras classificações de taxonomia: por assunto; por unidade; e funcional. Para Vital (2012), com base em Blackburn (2006), essas classificações podem ser descritas da seguinte forma:

- Taxonomia por assunto: faz uso de um vocabulário controlado. Exige que o usuário tenha conhecimento do domínio explorado. A estrutura desta taxonomia é apresentada pela relação de termos genéricos e específicos;

- Taxonomia por unidade: está orientada pelo negócio da organização. Neste tipo de taxonomia, os usuários já estão familiarizados ao ambiente. De acordo com a

própria classificação, sua estrutura é apresentada pela relação das unidades de negócios.

- Taxonomia funcional é aquela que está organizada pela dinâmica funcional desenvolvida pela organização. Aqui, os usuários podem navegar pelo ambiente de acordo com as suas atividades; e a relação dessas atividades com outras, desempenhadas por outros usuários. Pode-se citar como exemplo a construção de *workflows*, em que o fluxo de atividades permeia duas ou mais cadeias de valor de uma empresa.

Em suma, de acordo com Vital (2012, p. 38), taxonomia pode ser entendida como um modelo de representação que visa atender necessidades informacionais de ambientes web e corporativos. Essa ideia pode ser corroborada na reunião das operações vistas nas taxonomias classificadas acima.

Considerando as contribuições do aperfeiçoamento digital frente ao desenvolvimento de taxonomias, Steimer e Luz (2015, p. 3) veem o sistema em dois tipos: taxonomia *front end* e taxonomia *back end*. A diferenciação entre esses dois tipos baseia-se no fato de que: taxonomia *front end*, ou de navegação, contempla Arquitetura da Informação, usabilidade e *search engine optimization*; e uma taxonomia *back end*, ou estrutural, pode ter por base metadados compartilhados por uma organização, tal como a descoberta de conhecimento via processos de Gestão do Conhecimento. O estudo publicado pelos autores considerou “[...] áreas relacionadas à navegação por Arquitetura da Informação e Gestão de Conteúdo em ambientes digitais”. Assim, concluíram que “[...] a tendência é o desenvolvimento destes estudos, juntamente com a classificação por automação e aprendizado de máquinas”.

A discussão sobre esses tipos de taxonomia, tal como desses conceitos tecnológicos, implica no processo de atuação de profissionais da informação, pois, com o desenvolvimento tecnológico, esse escopo profissional de Bibliotecários, Gestores de Informação, Arquivistas e Documentalistas ampliou-se.

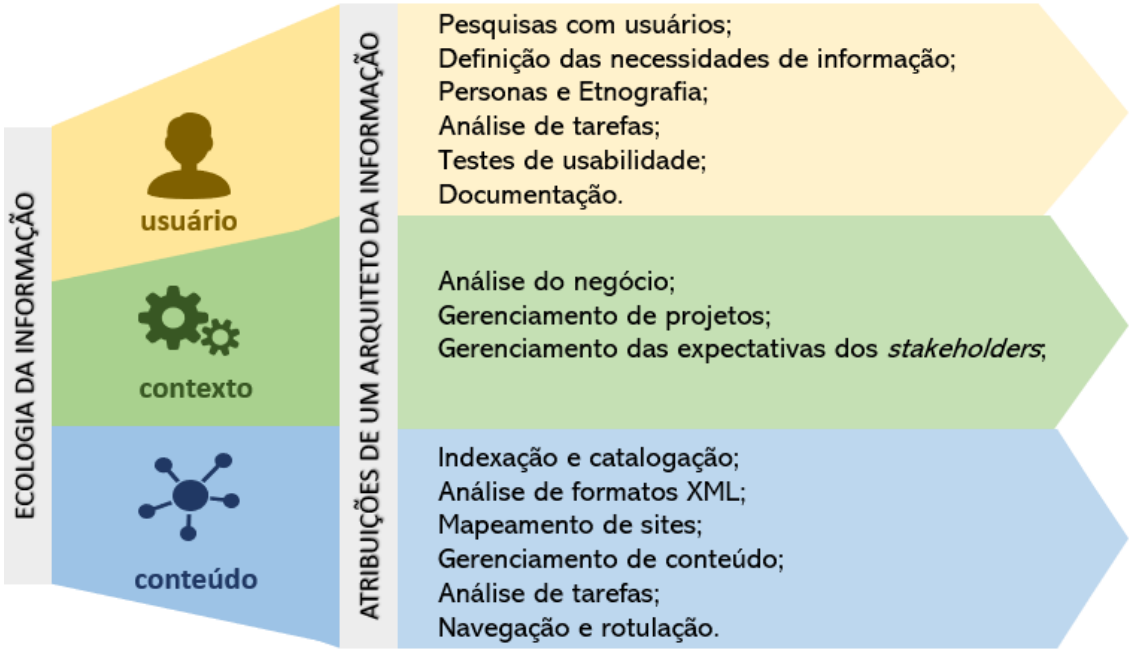
> Hoje profissionais da informação já atuam em organizações privadas, juntamente como área-meio para azeitar processos de Gestão do Conhecimento, inteligência competitiva e até mesmo pesquisa de mercado, para citar alguns exemplos. O desenvolvimento de uma taxonomia, tanto de *front* como de *back end*, pode ser entendida como curadoria de metadados e o profissional da informação tem os conhecimentos técnicos e conceituais necessários para este trabalho em específico. (STEIMER; LUZ, 2015, p. 8)

A ideia da Arquitetura da Informação (AI) é uma concepção funcional trabalhada em demasia nos estudos do publicitário e arquivista Charlley Luz, como os dois estudos citados (STEIMER; LUZ, 2015; LUZ, 2018). Aqui, pode-se trazer à tona a necessidade de profissionais da informação que detenham conhecimento sobre AI. Essa demanda por ser ocasionada pelo aprimoramento da tecnologia digital. Logo, para além de conceitos e processos, implica-se, também, que esse aprimoramento vem refletindo na formação desses profissionais.

No sentido de estreitar essa discussão sobre taxonomia, AI e tecnologia, de acordo com a matéria “Arquitetura da Informação e seu propósito” de Santana (2017, *online*), podem-se apontar três gerações da Arquitetura de Informação: a primeira, em 1970 e 1980, que teve como principais pontos: “[...] esclarecimento da necessidade de uma abordagem arquitetural e as analogias com arquitetura de construção”; a segunda, em 1990, “[...] teve como foco em sistemas web como conjuntos integrados de componentes de organizações individuais”; e a terceira geração, depois de 1990 e 2000. Esta última deve-se à “[...] informação como recurso corporativo com ferramentas de apoio de tecnologia da informação e técnicas”, que, conforme Santana, traz os principais pontos: “[...] definição explícita de princípios e teoria básica; desenvolvimento de arquiteturas multidimensionais; customização de frameworks de informação para as necessidades de organizações individuais; padrões e mapas de informação genérica”.

Profissionais da Informação podem atuar, por exemplo, como Arquitetos de Informação. Mas, quais são as atribuições deste tipo de profissional? O esquema abaixo foi construído para ilustrar o que Santana (2017) apresentou sobre a visão da ecologia da informação, reunindo os três elementos: usuários, como operadores da informação; contexto, como políticas, cultura, objetivos do negócio etc.; e, conteúdo, como os documentos, imagens, mídias etc.

Figura 84 - Ecologia da Informação e o Arquiteto de Informação



Fonte: O autor (2022) baseado em Santana (2017).

A AI, portanto, tem "[...] como objetivo influenciar positivamente na experiência do usuário para que ele busque e encontre as informações necessárias para executar sua tarefa ou aprender algo novo". Assim como os sistemas taxonômicos (e como propriedade desses sistemas), AI é rica em teoria e prática. Cabe, portanto, aos profissionais da informação, através de estudos dessas teorias e práticas: "[...] saber identificar as necessidades dos clientes; elaborar projetos para desenvolvimento de soluções; desenvolver *wireframes* e/ou protótipos; aplicar testes de usabilidades; e implementa soluções" (SANTANA, 2017, *online*).

Com base no esquema visto acima, e corroborando com Jacintho e Moreiro González, é necessário reconhecer que as atribuições mapeadas, por exemplo, refletem um novo contexto laboral e social em que profissionais da informação podem desenvolver habilidades e competências em suas funções. Para os autores, "[...] as empresas, indústrias e organizações, impulsionadas pela globalização, estão colaborando para a formação de novos perfis profissionais no mercado de trabalho" (2019, p. 436). Logo, é preciso perceber também que esses perfis precisam ser desenvolvidos na universidade e em outros espaços de aprendizagem, como os portais educativos.

De acordo com os extratos analisados, assim como o sistema tesauro, as taxonomias foram encontradas em diferentes operações sociais, culturais etc., prevalecendo a ideia lógica aplicada em portais na web. Essas percepções estão apresentadas na figura 85, a seguir.

Figura 85 - Exemplos de taxonomia

Fonte: O autor (2022) baseado em dados da pesquisa.

Ao observar o quadro acima, de acordo com as taxonomias apresentadas pelos autores, corrobora-se a ideia de que este tipo de SOC é aplicado em distintos domínios do conhecimento, como: portais corporativos; saúde; cultura; gestão estratégica da informação e do conhecimento; auditoria etc. Ademais, é possível perceber aqui exemplos de taxonomias orientadas à era do Big Data (discutida na seção 4.1), por relacionar a aplicação deste sistema às técnicas de aprendizado de máquinas e mineração de textos, como os trabalhos de Silva, Freitas e Silva (2014), Santos, Martins e Kaneoya (2016) e Braga (2016).

Os processos que originam taxonomias são orientados por métodos distintos. Tal como visto na seção anterior (sobre tesauros), existem as Normas reguladoras para guiar profissionais de informação à construção desse sistema; mas, ainda

assim, existem as operações adicionadas nos processos adotados para construir taxonomias.

Moreiro González (2011), ao tratar dos processos para construção de taxonomias, elencou princípios norteadores, como:

- **Mineração de dados**: trata-se da identificação dos objetos que serão incluídos na estrutura taxonômica. Essa identificação pode ser realizada por meio de entrevistas com especialistas, documentos publicados e instrumentos classificatórios ou terminológicos. Para este último pode-se citar como exemplo o tesauro;

- **Análise dos documentos/informações que foram agregadas à taxonomia**: por se tratar de um sistema que objetiva a navegação de pessoas usuárias em portais na web, é necessária a construção de uma superestrutura classificatória, que representa o corpus de documentos analisados e agregados ao sistema;

- **Elaboração da superestrutura classificatória da taxonomia:** neste ponto, cabe a reflexão de todas as questões que deverão ser atendidas na taxonomia, como: quem são as pessoas usuárias? Quais setores deve atender? Quais os processos e tarefas de gestão? Para que, em seguida, determine-se qual o melhor método para construí-la; quem serão os responsáveis por sua construção; quais os princípios norteadores? Quais as formas (modelos) de representação do sistema? Qual o melhor software para desenvolvimento deste sistema?

Aqui, vale ressaltar que “[...] toda taxonomia é fruto de um processo representativo e classificatório, por isso é produto de uma construção representativa do estado e da visão do conhecimento de seus elaboradores” (MOREIRO GONZÁLEZ, 2011, p. 62).

- **Verificação por especialista**: nesta etapa, devem-se cumprir duas funções: confirmar o trabalho realizado e transferir o conhecimento do processo de realização.

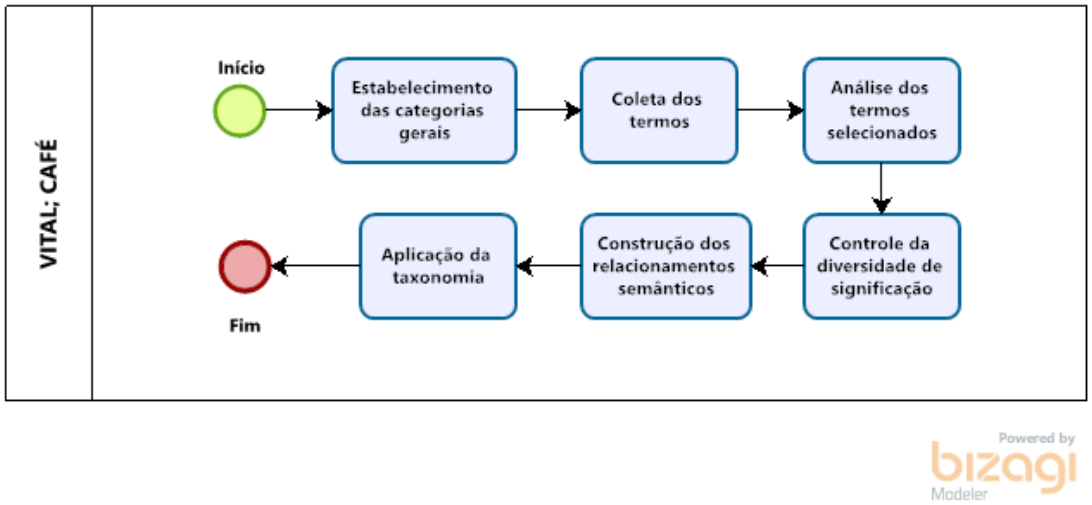
Ao tratar da transferência de conhecimento do processo de construção desse sistema, pode-se ter como exemplos os mapeamentos dos processos para desenvolvimento de taxonomias. Como um dos objetivos desta tese, e de acordo com os extratos consultados, esses mapeamentos estão apresentados abaixo.

As pesquisas de Aganette e Teixeira (2017), Bernardino, Simões, Gracioso e Costa (2015), Cunha e Araújo Júnior (2017), Plaza-Carvajal, Dutra e Macedo (2017),

Santos, Martins e Kaneyoa (2016), Vital e Café (2011) e Wu e Yang (2015) foram detalhadas a seguir e mostraram semelhanças no processo de construção dos sistemas propostos.

Vital e Café (2011) modelaram o processo de construção de taxonomias orientadas aos ambientes corporativos. Para as autoras o desenvolvimento desse sistema está orientado por cinco etapas. A última etapa do diagrama abaixo indica a aplicação da taxonomia em ambiente digital.

Figura 86 - Mapeamento dos processos para construção da taxonomia por Vital e Café



Fonte: baseado em Vital e Café (2011).

Para a primeira tarefa (estabelecimento das categorias gerais), de acordo com Vital e Café (2011), deve-se seguir duas abordagens: dedutiva e indutiva. Pode-se apontar como abordagem dedutiva aquela que reúne múltiplos especialistas do domínio que se quer desenhar e, por meio de discussões, desenha-se um escopo, elencando os principais pontos discutidos. Por outro lado, a abordagem indutiva seria exemplificada por meio da consulta aos materiais, manuais e guias de instruções, realizada por pessoas responsáveis pelo desenvolvimento do sistema. Esse último método, portanto, pode ser discutido, também, por um grupo de especialistas. As autoras, para tanto, ressaltam sobre a "[...] necessidade de denominar corretamente as categorias, notando aspectos de variação conceitual e terminológica" (VITAL; CAFÉ, 2011, p. 46).

Na tarefa de coleta de termos, as autoras defendem a ideia do carecimento de se consultar materiais disponíveis que tratem do tema e reconhecem que os aspectos institucionais devem servir como guia para desenvolvimento desse

sistema. Nesse sentido, a etapa seguinte à coleta de termos é a identificação da padronização desses, para que o sistema tenha uma comunicação efetiva com o usuário e, posteriormente, o controle da diversidade de significados.

Na sequência das tarefas acima, parte-se para a construção dos relacionamentos semânticos. Daí, "[...] leva-se em consideração as relações semânticas entre os termos, que podem ser de equivalência, hierárquicas e não-hierárquica, assim como polihierarquias" (VITAL; CAFÉ, 2011).

Como resultado, as autoras discutem sobre duas visões. A primeira, sobre Representação da Informação:

> Em relação à aplicação da taxonomia na representação da informação (classificação ou indexação), o desenvolvimento das etapas anteriores influencia diretamente e, em consequência, afeta significativamente a recuperação da informação. Termos coletados e analisados, assim como relações semânticas estabelecidas, possibilitarão ou não a indexação adequada. Uma taxonomia bem construída facilitará atingir princípios e critérios de tratamento da informação determinados pelo portal corporativo. (VITAL; CAFÉ, 2011, p. 51)

A segunda, acerca da aplicação da taxonomia para Organização da Informação:

> Em relação à organização da informação recuperada do acervo, a taxonomia proporciona a categorização dos resultados de uma busca, facilitando a identificação das relações entre os assuntos. O mecanismo de busca do portal deve apresentar a saída dos resultados organizados nas categorias previamente definidas. Dessa forma, o usuário perceberá outros assuntos relacionados ao tema procurado que, muitas vezes, não lhe haviam ocorrido. As categorias gerais estabelecidas na etapa 'A' para a organização dos termos são as mesmas utilizadas no momento da recuperação. É necessária a conexão entre a documentação do portal e as categorias da taxonomia no momento da aplicação, para que a taxonomia auxilie no processo de recuperação, com link direto aos recursos informacionais. (VITAL; CAFÉ, 2011, p. 52)

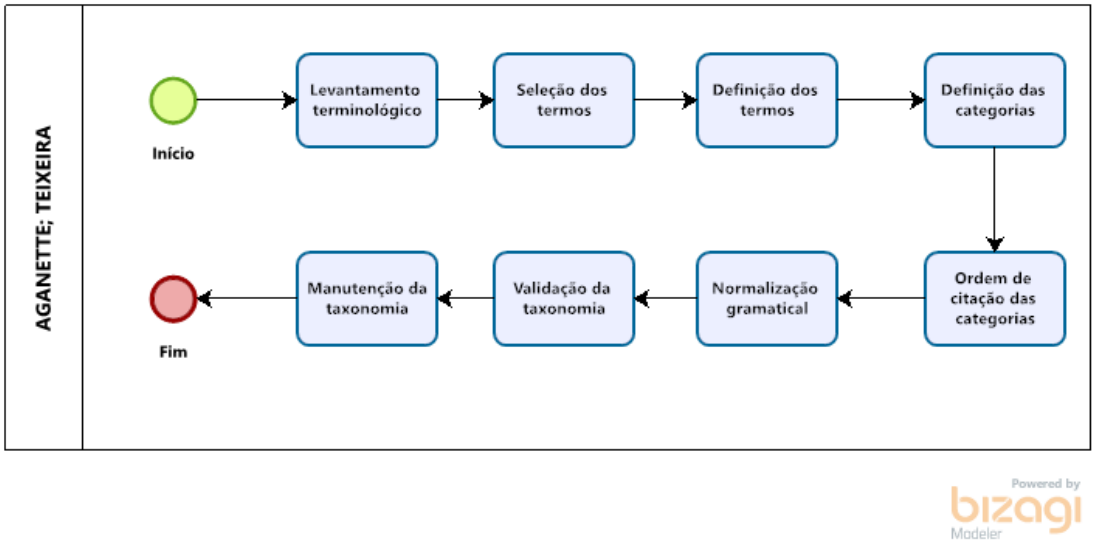
O estudo dessa modelagem buscou elucidar as tarefas básicas para a construção de taxonomias em portais corporativos, que, apesar de indicarem taxonomias como a estrutura central de organização e representação da informação, a literatura da CI não apresenta indicações metodológicas claras.

Tão logo, faz-se pertinente observar a modelagem dos diagramas aqui propostos, facilitada por aplicativos digitais, pois ela permite que essa tese contribua para a compreensão dos processos adotados por autores e pesquisas que buscaram o desenvolvimento e aplicação desse sistema. E que a reunião desses

mapeamentos possa fornecer “[...] o devido aprofundamento teórico e metodológico necessário nas etapas da construção” (VITAL; CAFÉ, 2011, p. 52).

Nessa mesma discussão sobre desenvolvimento de taxonomias corporativas, o estudo de Aganette e Teixeira (2017) buscou propor um procedimento operacional que orientasse, de maneira detalhada, as etapas de construção e implementação de taxonomias corporativas em sistemas de recuperação da informação. As autoras exploraram como guia a Norma ANSI/NISO Z39. 19 de 2005, mas enfatizaram que essa “[...] se trata de procedimentos facultativos” (AGANETTE; TEIXEIRA, 2017, p. 11). O diagrama abaixo ilustra o percurso metodológico orientado pelas autoras.

Figura 87 - Mapeamento dos processos para construção da taxonomia por Aganette e Teixeira



Fonte: O autor (2022) baseado em Aganette e Teixeira (2017).

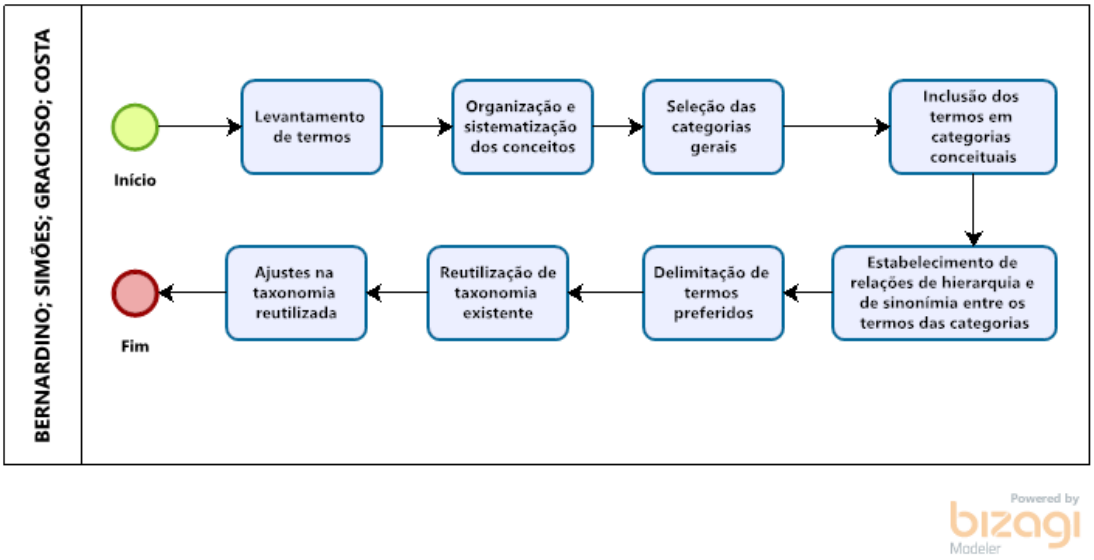
Tal como Aganette e Teixeira (2017) e Vital e Café (2011) defenderam a ideia da escassez de pesquisas que apresentem a modelagem de procedimentos operacionais para desenvolvimento de taxonomia. A idealização desta tese traz consigo discussões que refletem essas considerações, pois, aqui, reúne-se não apenas um modelo operacional guiado para um determinado domínio, mas uma pluralidade de aplicações, métodos, tarefas e ações em um fluxo operacional ilustrado e discutido, tornando essa pesquisa uma fonte rica de sistematizações que podem ser adotadas como modelos aplicáveis para construção de taxonomias.

O modelo operacional apresentado por Aganette e Teixeira (2011) objetivou, portanto, a definição das atividades necessárias para aplicação de taxonomias em

ambientes corporativos, garantindo a qualidade, estruturação e recuperação das informações. Conforme as autoras, ele é aplicável em qualquer área, setores e processos de uma organização. Para efetivar seu desenvolvimento e aplicação, as autoras enfatizam que “[...] é de responsabilidade do Gestor da Informação as atividades descritas neste procedimento” (AGANETTE; TEIXEIRA, 2017, p. 14).

Além do desenvolvimento de taxonomias para ambientes corporativos, foram explorados outros campos de aplicação deste sistema, como o estudo de Bernardino, Simões, Gracioso e Costa, que objetivou descrever os resultados de pesquisas outrora desenvolvidas: “[...] a elaboração de uma taxonomia a ser utilizada para representação e recuperação de conteúdos sobre bens culturais rurais, em software criado para este fim” (BERNARDINO; SIMÕES; GRACIOSO; COSTA, 2015, p. 382). O percurso metodológico seguido pelas autoras está apresentado a seguir.

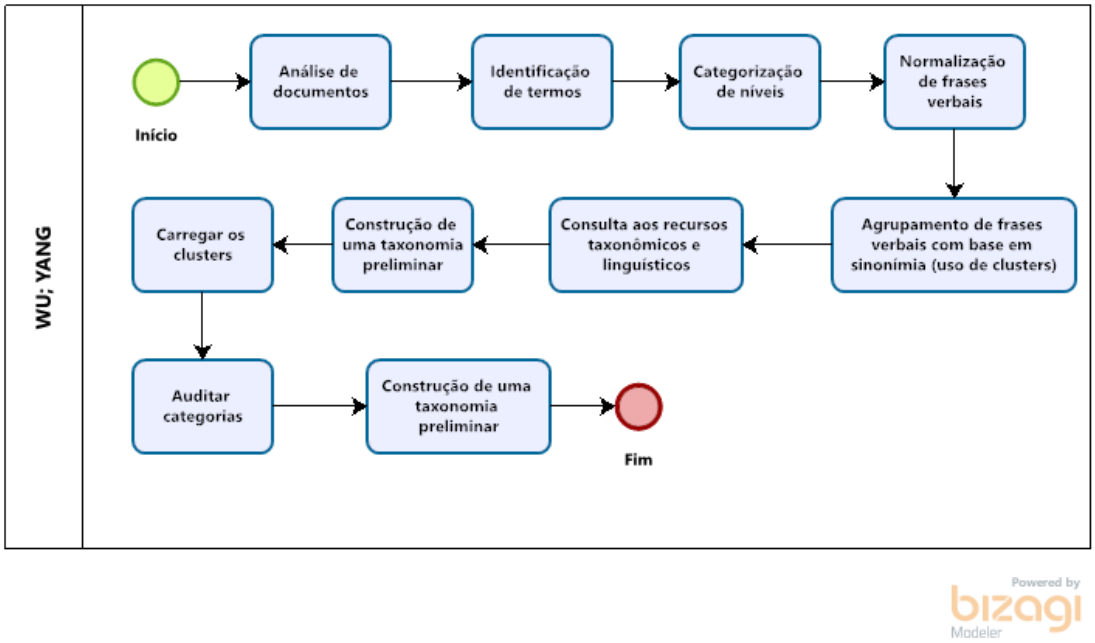
Figura 88 - Figura 88 –Mapeamento dos processos para construção da taxonomia por Bernardino, Simões, Gracioso e Costa



Fonte: O autor (2022) baseado em Bernardino, Simões, Gracioso e Costa (2015).

Tal como o interesse do uso de taxonomia para evidenciar uma representação de cunho social e histórico, o estudo de Wu e Yang objetivou a fundamentação lógica, metodológica e procedimental do desenvolvimento de taxonomias no domínio do derramamento de óleo como apoio à descoberta de conhecimentos. Os procedimentos seguidos estão mapeados no diagrama a seguir.

Figura 89 - Mapeamento dos processos para construção da taxonomia por Wu e Yang

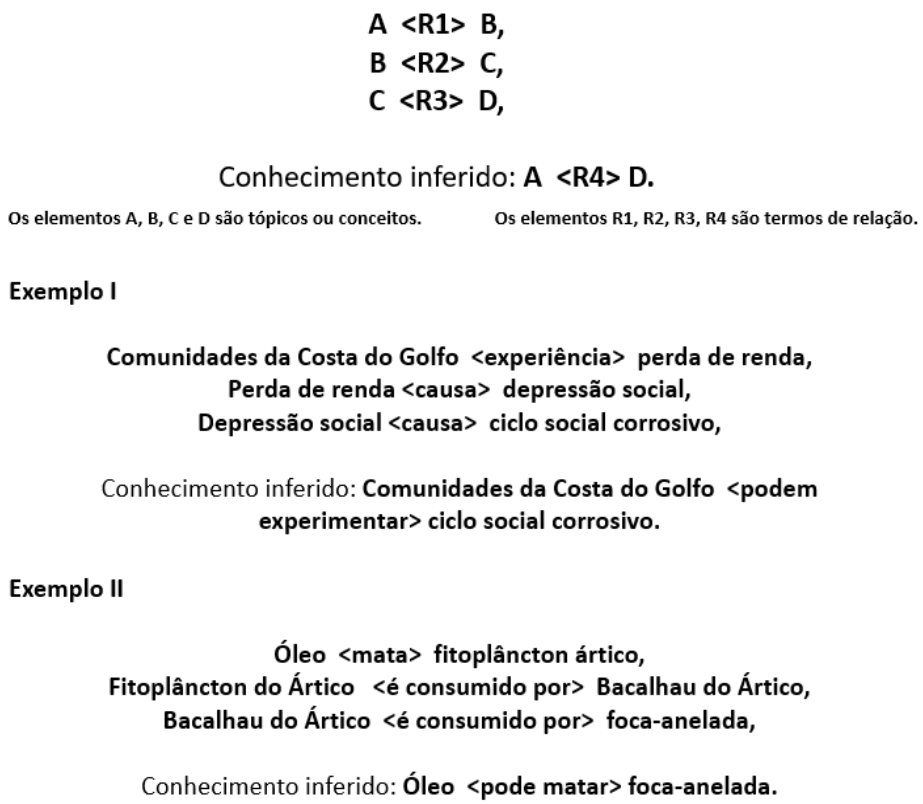


Fonte: O autor (2022) baseado em Wu e Yang (2015).

Os autores potencializaram o uso de taxonomias de relações semânticas no domínio do derramamento de óleo. Para tanto, tiveram como apoio dois eixos: o primeiro, sobre o incidente de derramamento de petróleo no Golfo do México em 2010; como elemento norteador (segundo eixo), o uso de mapa de tópicos de derramamento de óleo, criado para entender o impacto desta ação. Segundo Wu e Yang (2015, p. 116): “[...] as relações semânticas têm muitas aplicações na recuperação de informação, respostas a perguntas e organização do conhecimento” e, também, “[...] facilitam a descoberta de conhecimentos por meio da inferência”.

Um esquema lógico do uso de taxonomia da relação do derramamento de óleo, para efetuar inferências, é descrito a seguir:

Figura 90 - Esquema lógico do uso de taxonomia da relação do derramamento de óleo para efetuar inferências



Fonte: O autor (2022) adaptado de Wu e Yang (2015, p. 115).

Diferente das taxonomias vistas até aqui, em termos conceituais, processuais e operacionais, o uso da taxonomia acima dispensa critérios para navegabilidade. Ademais, nota-se que, no diagrama que descreve o fluxo processual, adotado para desenvolvimento da taxonomia, o uso de um conceito adotado pela Ciência da Computação (*cluster)* é apresentado como uma das tarefas.

Vale salientar que *cluster*, para os informáticos, trata-se da "[...] conexão entre dois ou mais computadores com o propósito de melhorar o desempenho dos sistemas na execução de diferentes tarefas" (SILVA, 2021, *online*).

A ideia da abordagem ao *cluster* é de unir e automatizar processos de análises dos termos ou dos objetos que o constituem. Tem-se aqui a aplicação de tecnologias para facilitar o processo de desenvolvimento de taxonomia. Os autores descreveram com precisão as etapas em que utilizaram *cluster*: após a normalização das frases verbais, podem-se agrupá-las com base na sinonímia dos termos; tão logo o processamento dessas frases seja efetuado por meio de *cluster*. Na pesquisa de Wu e Yang, por exemplo, 15 grandes *clusters* foram desenvolvidos para receber

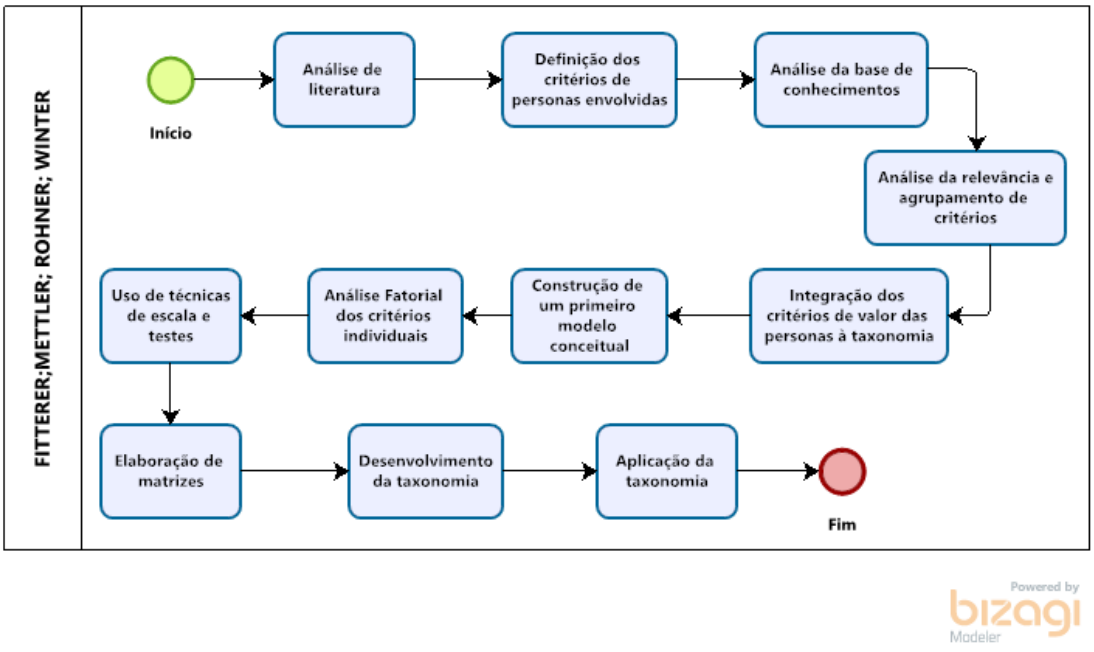
896 frases. Por fim, é necessário carregar esses *clusters*, um a um, na taxonomia que está sendo construída. Destaca-se, aqui, que este processo deve ser interativo.

Para um resultado mais assertivo, os autores utilizaram outros meios de garantia, como materiais de consultas, para somar qualidade à taxonomia, por exemplo: dicionários, outros recursos taxonômicos e linguísticos. Posteriormente, é preciso auditar as categorias e ajustá-las conforme as necessidades percebidas (WU; YANG, 2015).

Essa tarefa de explorar outros materiais de consultas, no desenvolvimento de taxonomias, pode ser percebido, também, na pesquisa de Fitterer, Mettler, Rohner e Winter (2011). Esses autores, portanto, após as etapas iniciais de estudos do domínio, inferem a tarefa de analisar bases de conhecimentos. Percebe-se, aqui, que a consulta a outros recursos, ao desenvolvimento de sistemas de organização do conhecimento, é essencial para garantir o sucesso do usuário nas tarefas de busca, recuperação, acesso e uso de informações.

O mapeamento dos processos do fluxo, para desdobramento da taxonomia, está apresentado a seguir.

Figura 91 - Mapeamento dos processos para construção da taxonomia por Fitterer, Mettler, Rohner e Winter



Fonte: baseado em Fitterer, Mettler, Rohner e Winter (2011).

O fluxo mapeado acima teve o objetivo de propor uma taxonomia para fortalecer o valor de Sistemas de Informação em Saúde por profissionais desse

domínio, como os da Enfermagem. Assim, “[...] uma taxonomia do valor de Sistemas de Informação em Saúde de uma perspectiva de aceitação de tecnologia é projetada como resultado da revisão da literatura” (FITTERER; METTLER; ROHNER; WINTER, 2011, p. 47).

As primeiras etapas são aquelas consideradas comuns, desde o estudo do ambiente aos critérios de seleção de termos candidatos ao sistema. Contudo, o estudo verticalizado, mais aprofundado, dos sujeitos do domínio, neste fluxo, diferencia-o dos demais vistos até aqui.

Os autores despenderam esforços significativos para compreender os valores individuais dos grupos analisados, no sentido de entender as razões da resistência ao uso de sistemas de informação em saúde. Em suma, a taxonomia “[...] se mostrou útil para integrar considerações estratégicas com capacidades técnicas/ funcionais baseadas nos requisitos das partes interessadas expressos pelos critérios da taxonomia” (FITTERER; METTLER; ROHNER; WINTER, 2011, p. 57).

A exploração de outros recursos, como consultas a materiais técnicos, dicionários e outros documentos textuais, é evidenciada em todos os fluxos aqui desenhados. Com o intuito de minimizar esforços manuais, por meio do auxílio de mecanismos digitais, pesquisas vêm sendo desenvolvidas com o objetivo de aprimorar determinadas etapas no fluxo de desenvolvimento de taxonomias. Como exemplo desse aperfeiçoamento, cita-se as práticas de automação, a exemplo da pesquisa de Silva, Freitas e Silva (2014).

A fim de compreender o processamento automático como um aperfeiçoamento metodológico, os autores consideram que “[...] em aplicações que envolvem o processamento de conteúdo textual, esse conhecimento de mundo pode significar, por exemplo, compreensão sobre o conteúdo de um texto”. Aqui, pode-se relacionar, por inferências para descobertas de conhecimento em um determinado documento. Isso porque “[...] taxonomias podem ser usadas para identificar em um texto original seus principais tópicos para subsidiar as mais diversas áreas”. (SILVA; FREITAS; SILVA, 2014, p. 55).

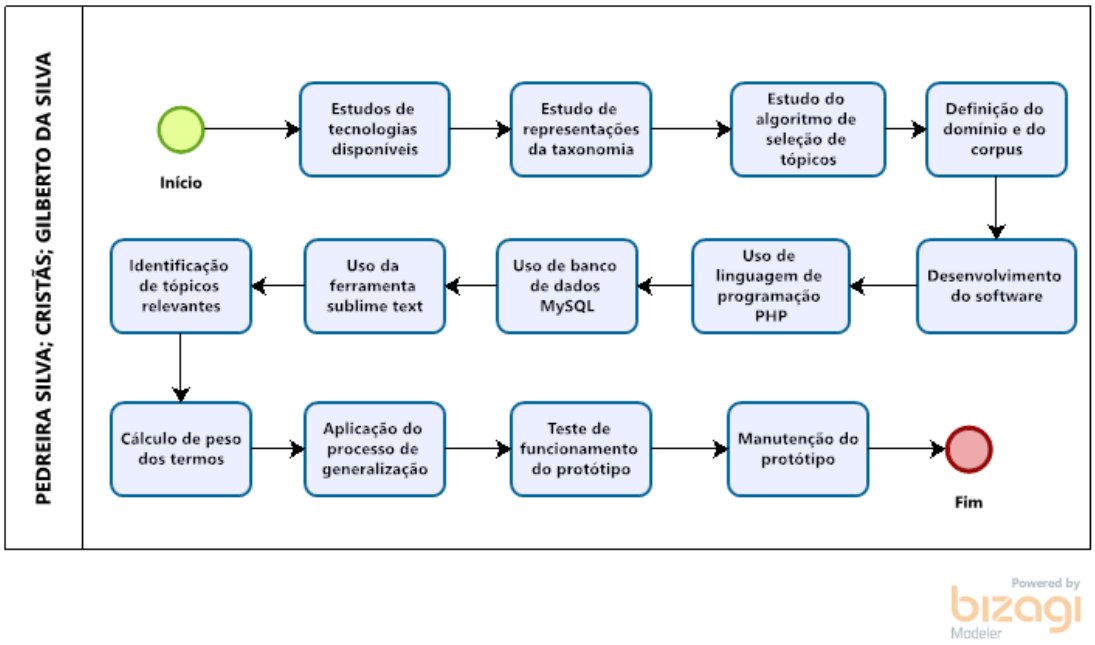
De acordo com as ideias acima discutidas, os autores apresentaram a pesquisa intitulada “*Desarrollo de un software para la detección automática de tópicos en documentos textuales basada en taxonomía”*, que objetivou, portanto, a adoção do “[...] uso de taxonomias como uma das estratégias para identificação de

tópicos e como suporte para o desenvolvimento de software que utilize este conhecimento". Assim, a ferramenta desenvolvida foi denominada EXTRATOP (Extração Automática de Tópicos), "[...] que combina características linguísticas e estatísticas para a detecção automática dos tópicos de um texto" (SILVA; FREITAS; SILVA, 2014, p. 56).

Em atenção ao objetivo da proposta acima, pode-se considerar que o uso automático de extração de tópicos, por meio de estruturas taxonômicas, facilita, também, as atividades inerentes ao domínio da organização e representação da informação, como na execução das atividades de indexação e catalogação.

As etapas para desenvolvimento da taxonomia estão apresentadas a seguir.

Figura 92 - Mapeamento dos processos para construção da taxonomia por Silva, Freitas e Silva



Fonte: O autor (2022) baseado em Silva, Freitas e Silva (2011).

Após as etapas do estudo do domínio, na figura acima trata-se das quatro primeiras tarefas, que foram importantes para dar início ao desenvolvimento do software. A ênfase dada a esse desenvolvimento consistiu na "[...] implementação de uma base de dados contendo informações de natureza taxonômica, o desenvolvimento do contorno do algoritmo de pontuação e a implementação de uma interface web". Linguagem de programação em PHP, Java Script e bancos de dados MySQL foram utilizados para desenvolvimento do software (SILVA; FREITAS; SILVA, 2014, p. 56).

Sobre a execução do software, os autores explicam que:

> [...] uma fonte de texto é fornecida como entrada para o sistema e para a detecção de tópicos, são realizadas as seguintes etapas:
> 1. Primeiro, o texto de origem é pré-processado, a pontuação tradicional (por exemplo, ponto final, ponto de exclamação e ponto de interrogação) sendo removida;
> 2. Os conceitos subjacentes ao texto são detectados, com base na taxonomia;
> 3. As palavras irrelevantes são eliminadas do texto e, em seguida, os tópicos são selecionados e classificados de acordo com sua proximidade com os conceitos de taxonomia;
> 4. Por fim, os tópicos que indicam o conteúdo principal do documento aparecem como resposta final. (SILVA; FREITAS; SILVA, 2014, p. 56)

Por uma avaliação geral dos resultados obtidos, os autores consideraram satisfatório. Contudo, perceberam que ajustes devem ser efetuados no sentido de melhorar o conteúdo da taxonomia. Esses ajustes têm como motivação as propriedades da Língua Portuguesa, como as ambiguidades. Em contrapartida, como uma perspectiva de solução, os autores visaram o rumo de desenvolvimento da web semântica. Para eles, portanto, "[...] a tendência é que o processamento semântico dos documentos se torne cada vez mais necessário, tornando recursos como taxonomias, por exemplo, cada vez mais importantes" (SILVA; FREITAS; SILVA, 2014, p. 60).

O fluxo acima é um dos exemplos que corroboram com as discussões realizadas sobre as fases dos Sistemas de Organização do Conhecimento, na seção 4.1 desta tese. A Era Big Data, como exemplo de uma dessas fases, potencializou o uso técnicas sofisticadas. Por isso, explorou-se a pesquisa intitulada "*Extração semiautomática de taxonomia para domínios especializados usando técnicas de mineração de textos*", de Braga (2016).

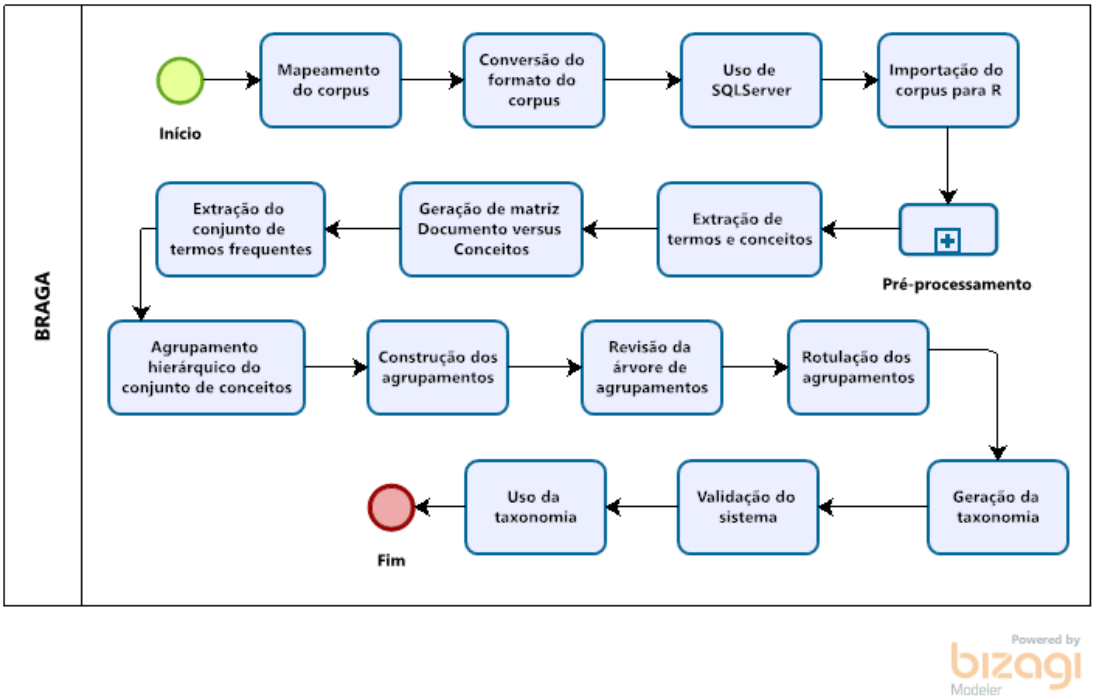
A pesquisa de Braga (2016, p. 177):

> [...] explorou o processo de descoberta de conhecimento e utilizou as técnicas de mineração de textos em uma base de dados textuais de artigos científicos no desenvolvimento de uma metodologia para a geração semiautomática de uma taxonomia de conceitos apresentando uma nova forma de organização desses conhecimentos.

As fases seguidas por Braga (2016) buscaram cobrir as etapas do processo de mineração de textos. Para tal, elas podem ser desenhadas sob quatros eixos principais: **geração de corpus** – contempla a extração de documentos e importação

para ambiente R; **pré-processamento** – compreende a remoção de *stopwords*, separadores, números, pontuação, etc., a extração dos conceitos e geração de matriz documentos *versus* conceitos; **processamento** – busca a extração dos conceitos frequentes, agrupamento hierárquico baseado nesses conceitos, construção da árvore e rotulação dos agrupamentos; **pós-processamento** – responsável pela geração da taxonomia e sua devida validação. A modelagem desse fluxo esta apresentada a seguir.

Figura 93 - Mapeamento dos processos para construção da taxonomia por Braga



Fonte: O autor (2022) baseado em Braga (2016).

A primeira fase de geração do corpus contou com o uso do banco SQLServer, e, posteriormente, importou-se o corpus para ambiente R. O *software open source* R é visto como uma "[...] poderosa ferramenta para análise de dados e representação gráfica que oferece excelente estrutura para fins de mineração de texto". Assim, "[...] essa estrutura permite aos usuários trabalhar de forma eficiente com textos e metadados correspondentes e transformar os textos em representações estruturadas" (BRAGA, 2016, p. 179).

A fase de pré-processamento consiste nas seguintes etapas (BRAGA, 2016): *tokenização* – é utilizada para decompor o documento em cada termo que o

compõe; *stopword* – remoção de dígitos e pontuações; tratamento do texto – eliminação de espaços em branco e conversão dos textos em minúsculas; e etiquetagem *Parth-of-Speech* (POS), isto é, etiquetagem gramatical. Ainda nesta fase, são realizadas as tarefas de extração de termos e conceitos, por uso de um tesauro, para auxiliar na consulta de terminologias; e geração de matriz documentos *versus* conceitos.

A fase de processamento constituiu-se no desenvolvimento do algoritmo para agrupamentos de documentos, da qual o Braga (2016) evidencia os estágios de: construção dos grupos iniciais, criação de uma árvore em tópicos e manutenção da árvore construída. Nela, também se inclui a tarefa de rotulagem dos grupos, que procedeu da seguinte forma:

> - para cada grupo gerado, o algoritmo extrai um conjunto formado pelos conceitos mais frequentes. Um conceito é considerado frequente se estiver contido em uma quantidade mínima de documentos desse grupo. O suporte é fornecido para definir a quantidade mínima de documentos;
> - depois de selecionar os conceitos mais frequentes, o algoritmo nos permite expandir os conceitos originais usando o dicionário de sinônimos ETDE/ INIS, somando-lhes os termos equivalentes ao conceito, isto é, aqueles com a notação UF no dicionário de sinônimos;
> - de acordo com o esquema de enriquecimento proposto associado a cada um desses conceitos, seus termos equivalentes no tesauro são extraídos. Por exemplo, se o conceito principal é “Nuclear Energy”, seu termo equivalente seria “Atomic Energy”; Se o conceito principal for “Radiation protection”, os termos equivalentes seriam “health physics”, “nuclear safety”, “protection (radiation),” “radiation hygiene”, “radiation safety”, “radiological protection” e “safety” (Nuclear). Assim, se a consulta do usuário contiver o termo “Atomic Energy”, os documentos que contenham esse termo serão recuperados assim como aqueles que contenham “Nuclear Energy”. Braga (2016, p. 181)

Na visão de Braga (2016), a metodologia proposta pelo autor, representada no fluxograma desenhado acima, apresentou um novo produto para descoberta de conhecimento em documentos, por meio da construção de sistemas como taxonomia.

> O conhecimento gerado pela taxonomia pode ser usado para facilitar processos de organização e recuperação de informações, bem como sua própria compreensão da coleção textual organizada, ou mesmo servir como suporte para sistemas de suporte à decisão. Mas é importante que as técnicas sejam desenvolvidas para auxiliar especialistas no domínio, a fim de facilitar a compreensão e o uso dos conhecimentos adquiridos. (BRAGA, 2016, p. 185)

O uso de técnicas de mineração de textos é percebido para além do desenvolvimento de taxonomias. Outro sistema, por exemplo, que se utiliza dessas

técnicas, é a ontologia. Ontologia é mais um dos sistemas alvo de estudo desta tese, e está discutido na seção seguinte.

Conforme percebido no quadro sobre as concepções conceituais e os fluxos mapeados dos processos metodológicos, compreende-se que o sistema taxonômico foi aprimorado e tem como resultado desse aprimoramento as múltiplas facetas tecnológicas digitais, nas condições de ferramenta e espaço de construção, aplicação, operação e uso. Assim, ele vem recebendo maior atenção nos últimos anos, devido à percepção da sua capacidade operacional que, de acordo com os autores visitados nos processos desenhados, pode estar orientado a objetivos, usuários e operações distintas, como: estratégias organizacionais internas; processos de *onboarding*; descoberta de conhecimento nas organizações; navegação em ambientes digitais; treinamento e desenvolvimento de pessoas; sistematização de conhecimento, servindo como apoio às tarefas de mineração de dados; entre outros.

Corrobora-se o discurso acima quando Simões *et al.* (2016) considera que, apesar de ser vista como um sistema de classificação, a taxonomia tem características próprias e distintas, como: a aplicação em ambientes digitais e em web semântica; o alto nível de organização; a estrutura vocabular; a flexibilidade e o dinamismo.

Estas, hoje, são possíveis devido ao desenvolvimento de outros mecanismos que auxiliam no avanço deste sistema, enriquecendo-o como uma estrutura flexível e transformando, progressivamente, em um potencial sistema de organização do conhecimento.

4.4 Ontologia: conceitos e processos

Na seção 3, que apresentou as discussões introdutórias dos sistemas de organização do conhecimento (SOC), discutiram-se exemplos de SOC e, entre eles, ontologia. Tal como outros sistemas, isto é, tesauro e taxonomia, o termo ontologia tem origem em outra Ciência: a Filosofia. Aplicada aos estudos relacionados à Metafísica, a fim de compreender como as coisas são e de que forma se comportam, o conceito recebeu atenção advinda de outros estudos científicos, mas de caráter não filosófico.

A Ciência da Computação percebeu, nas técnicas orientadas à recuperação de informação, um tipo de sistema que, além da arquitetura, abordasse propriedades sobre os termos a serem processados. Neste mesmo sentido, tal como objeto de estudo também, a Ciência da Informação percebe neste sistema uma de suas atividades para exploração dos exercícios de organização e representação da informação e do conhecimento.

De acordo com o quadro bibliográfico explorado, foram mapeadas as concepções conceituais e processuais apresentadas pelos autores devidamente referenciados. O quadro abaixo aborda o sistema ontologia.

Quadro 11 – Concepções conceituais sobre ontologia.

ANO	AUTOR(A)(ES)(AS)	CONCEPÇÕES CONCEITUAIS
2010	KORNYSHOVA; DENECKÈRE	O termo ontologia é retirado da filosofia, onde Ontologia é um relato sistemático de Existência.
2010	SOUZA; SILVA; SILVA; AZEVEDO	A palavra ontologia vem do grego ontos (ser) + logos (palavra); O termo ontologia pode ter vários significados dependendo da realidade em que esteja sendo aplicado, visto que é usado para descrever domínios e estes domínios podem abordar um determinado assunto muito específico da sua realidade. Nesse sentido, o termo ontologia foi utilizado inicialmente na filosofia, passando, contudo, a ser utilizado também em outras ciências. Por ser usado em diversas ciências, existem muitas definições para ontologia que, em alguns casos podem se complementar.

2011	RAUTENBERG; TODESCO; STEIL	o uso de ontologias se justifica, pois estas podem ser utilizadas como uma fonte de aprendizado, publicação e referência de conhecimento pertinente. Em outras palavras, como elas são, presumidamente, sempre o resultado de um grande consenso sobre a estrutura de um domínio, as ontologias podem prover informações contextualizadas àqueles que querem aprender mais sobre o domínio. Simultaneamente, especialistas de domínio também podem usar ontologias para compartilhar seu entendimento da conceitualização e estrutura do domínio
2011	SCHIESSL; BRASCHER	Ontologia como explicação sistemática da própria existência se estende para o domínio de sistemas de informação (SI). De tal maneira, pode-se descrever uma ontologia pela definição de um grupo de termos representativos de determinado domínio. A tarefa desse domínio é representar formalmente essa existência. Nesse contexto, ontologia é um modelo que reflete uma visão de mundo. Depreende-se das argumentações anteriores que o objetivo dela é delimitar a quantidade de interpretações e caracterizar os significados básicos de categorias utilizadas para descrever determinado domínio. Logo, pragmaticamente, tem-se uma representação limitada e incompleta, porém útil, do domínio de interesse.; recentemente, as ontologias são reconhecidas como componente importante para a construção da Web Semântica; Nela, a informação está bem definida, o que possibilita o trabalho cooperativo entre computadores e pessoas. Para tanto, máquinas devem ter acesso a conjuntos de informações estruturadas e a grupos de regras de inferência que podem conduzir ao raciocínio automatizado

2012	YAMAOKA	Na computação e na ciência da informação, ontologia é um artefato elaborado para modelar o conhecimento de algum domínio ou assunto e permite especificar um vocabulário de asserções que pode ser utilizado por software.
2013	ALMEIDA, MENDONÇA, AGANETTE	Ontologia é um termo originalmente criado na Filosofia, mas que desde a década de 1990 tem aparecido também com frequência na literatura da Ciência da Computação e da Ciência da Informação; ontologias não são linguagens documentárias; ontologias são sistemas de organização do conhecimento; serve para definir um instrumento de organização; Essa configuração do instrumento, permite então que ele atenda ao requisitos de sistemas formais lógicos, e por consequência, do moderna tecnologia da informação; ontologias, mesmo que tenham grande apelo em função da possibilidade de automação e de uso na Web, na verdade são capazes de representar apenas uma pequena parcela do rico cenário abordado pela Ciência da Informação.
2013	ALMEIDA *et al*	Ontologia é um termo com origens na Filosofia, onde diz respeito ao estudo do que existe. Em Ciência da Informação tem sido usado para descrever um vocabulário formal composto por termos e relações organizados de acordo com princípios da ontologia filosófica.

2013; 2017	FERNEDA E DIAS	ontologia como um conjunto de conceitos interligados por relações semânticas; ontologias se apresentam como um novo instrumento a ser incorporado ao arsenal teórico e prático da Ciência da Informação. A aprendizagem de novos conceitos e novos recursos oferecidos pelas ontologias é um desafio para os profissionais da informação, mas que pode ser facilmente enfrentado utilizando todo o conhecimento teórico e prático acumulado durante a história da Ciência da Informação.
2013	HERRE	A importância das ontologias é cada vez mais reconhecida em campos diversos como e-commerce, web semântica, empresa, integração da informação, ciência da informação, modelagem qualitativa de sistemas físicos, processamento de linguagem natural, engenharia do conhecimento e bancos de dados. Ontologias fornecem especificações formais e definições harmonizadas de conceitos usados para representar o conhecimento de domínios específicos. Uma ontologia fornece uma estrutura unificadora para a comunicação, estabelece uma base para a organização e representação do conhecimento e contribui para a formação de teoria e modelagem de um domínio específico.
2013	SANTOS; CORREA; LAPA	A ontologia é para Ciência da Informação um sistema de organização do conhecimento que muito pode contribuir para a representação do conhecimento efetiva no ambiente digital. Porém são necessários mais estudos para se consolidar na Ciência da Informação por meio do mapeamento das pesquisas para que seja possível compreender o que é esse sistema e como pode ser utilizado de forma eficaz.

2014	MENDONÇA; ALMEIDA	Usadas como instrumentos para modelização do conhecimento de domínios diversos, as ontologias apresentam uma natural proximidade com técnicas consagradas pela Ciência da Informação.
2014	MOREIRA; SANTOS NETO	Na interface teórico-conceitual da ciência da informação com a ciência da computação, as ontologias são consideradas como artefatos desenhados especificamente para possibilitar a representação computacional de algum fenômeno. As ontologias são, portanto, modelos de algum aspecto da realidade, nos quais os conceitos são definidos de modo lógico-semântico em termos de seus limites e de suas correlações.
2014	PIETERSE; KOURIE	Uma ontologia é uma coleção armazenada eletronicamente que compreende um tesauro combinado com um conjunto de regras de inferência.; nosso conceito de ontologia se correlaciona com seu significado formal conforme usado na ciência da computação.; definimos uma ontologia como uma extensão de um tesauro. Ele contém itens que representam conceitos, seus atributos e relações em uma estrutura mais formal do que o exigido para tesauros em geral. Existem dois aspectos que distinguem uma ontologia de outros KOSs, a saber, o requisito de que a representação deve ser baseada em um formalismo e que deve incluir regras de inferência.

2014	RAMALHO; ZAFALON; OUCHI	as ontologias figuram como uma nova abordagem computacional de representação, descrição e organização de conteúdos informacionais, favorecendo a aproximação dos modelos clássicos de representação, utilizados tradicionalmente no âmbito da área de Ciência da Informação, com as novas tecnologias semânticas desenvolvidas ao longo dos últimos anos. As ontologias potencializam a representação dos relacionamentos existentes entre os conceitos que traduzem os conteúdos informacionais de modo a possibilitar representações semânticas que não poderiam ser obtidas a partir de descrições textuais, caracterizando-se como um tipo de fonte secundária, instituindo novas formas de acesso e favorecendo maior flexibilidade na recuperação de informações, por meio da realização de inferências.
2015	CAFÉ; SANTOS; BARROS	A ontologia enquanto campo de estudo tem sua origem na Filosofia, voltando-se às questões relacionadas à existência dos seres. No campo da organização da informação e como instrumento de representação do conhecimento, as ontologias são particularmente estudadas pela Ciência da Informação e Ciência da Computação. Esses estudos situam-se numa linha de interdisciplinaridade com outras ciências como a Lógica, Inteligência Artificial, Filosofia e a Linguística.
2016	SILVA; SOUZA	O conceito envolvendo “ontologia” tornou-se de interesse multidisciplinar para propósitos epistemológicos específicos na filosofia, linguística, biblioteconomia, inteligência artificial, ciência da Web, engenharia de software, lógica, apenas para citar alguns exemplos.; “ontologias” são de interesse como instrumento de organização de conhecimento que versa especialmente na análise de conceitos e relacionamentos de um domínio.

2016	PANSANI JUNIOR; FERNEDA	O termo ontologia possui definições distintas dependendo da área de conhecimento ao qual é citado; uma ontologia pode ser considerada como um vocabulário de representação, geralmente especializado em um domínio ou assunto, qualificado por conceituações de tipos de objetos e suas relações no mundo.
2017	MENDES NETTO; LIMA	A ontologia pode ser comparada a um esquema conceitual para um sistema, que fornecerá uma descrição lógica para compartilhamento dos dados. Na organização do conhecimento, uma ontologia fornece a representação de um domínio, possibilitando a criação de uma base de conhecimento que incluirá o conhecimento simbolicamente representado para resolver problemas específicos ou responder consultas sobre o domínio.
2017	SOUZA; ALMEIDA; SOUZA	Nessa perspectiva de organização do conhecimento, as ontologias têm se tornado uma alternativa para especificar, em linguagem formal, os conceitos e suas relações de uma determinada área de domínio;
2018	VILELA; GOTTSCHALG	As ontologias permitem representar o conhecimento de forma mais qualificada e completa, mais próxima da estrutura da mente humana; essa condição, reforçada pelo “apego à realidade” das ontologias, decorrente das suas raízes na Filosofia, permite também que as ontologias sejam reusadas em diferentes domínios do conhecimento, independentemente da língua em que foram elaboradas.

2018	MUCHERONI; SILVA; SANTOS	Com origens de natureza filosófica, as ontologias podem ter diferentes significados, dependendo da área de abordagem. Na Ciência da Informação e na Ciência da Computação, cujas atuações se entrelaçam cada vez mais, as ontologias são os instrumentos que permitem a nomeação formal e definição de categorias, propriedades e relações entre conceitos, entidades e dados atrelados a um ou mais domínios ou objetos de informação. Em uma correlação à Ciência da Informação tradicionalista, as ontologias podem ser comparadas, grosso modo, como uma releitura da indexação (e toda a documentação que a fundamenta), voltada para o universo semântico digital.
2019	SILVA; RIBEIRO; RODRIGUES	Pode auxiliar na tarefa de recuperação da informação em ambientes virtuais, visto que o excesso de informação disponível sem filtragem e sem tratamento acarreta grandes dificuldades no processo de busca e de recuperação da informação pelos usuários.

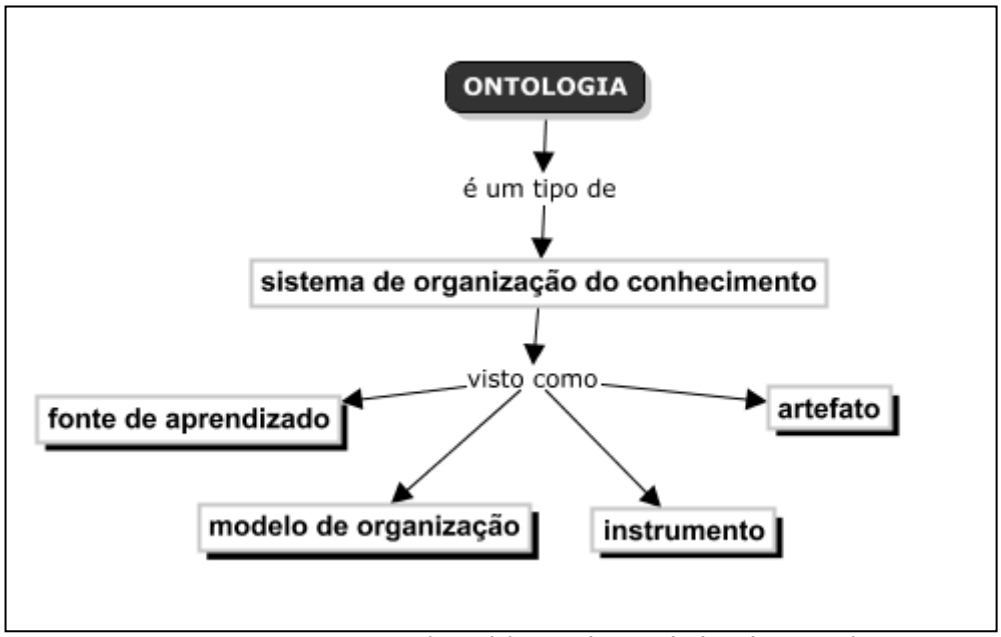
Fonte: O autor (2022) baseado em dados da pesquisa.

De acordo com as concepções acima, extraídas do quadro bibliográfico consultado para construção desta tese, percebe-se que a origem do conceito do termo “ontologia” é consenso entre os autores referenciados. Apesar da comunhão acerca do conceito primário, outros autores divergem quanto a percepção conceitual aplicada pela Ciência da Computação e pela Ciência da Informação.

Yamaoka (2012), conforme visto no quadro acima, reuniu um único conceito para ontologias, sob a ótica das duas ciências. Café, Santos e Barros (2015) e Lucas da Silva e Souza (2016), por outro lado, já percebem que o conceito é explorado por diversos outros campos, atrelados e não atrelados, à Ciência da Informação e à Ciência da Computação, tais como: Organização e Representação da Informação e do Conhecimento; Lógica; Inteligência Artificial; Filosofia; Linguística; Biblioteconomia; Engenharia de Software; entre outros. Em geral, para a Organização do Conhecimento, é comum encontrar o conceito deste sistema na mesma concepção de Souza, Almeida e Souza (2017).

Assim como os demais sistemas alvo de discussão, destaca-se que não é pretensão desta tese limitar as discussões sobre o conceito de ontologia, mas sim o de clarear concepções conceituais sobre este sistema e mapeá-lo quanto aos conceitos e processos aprimorados nos últimos anos. Logo, o mapa conceitual a seguir reúne as ideias sobre ontologia.

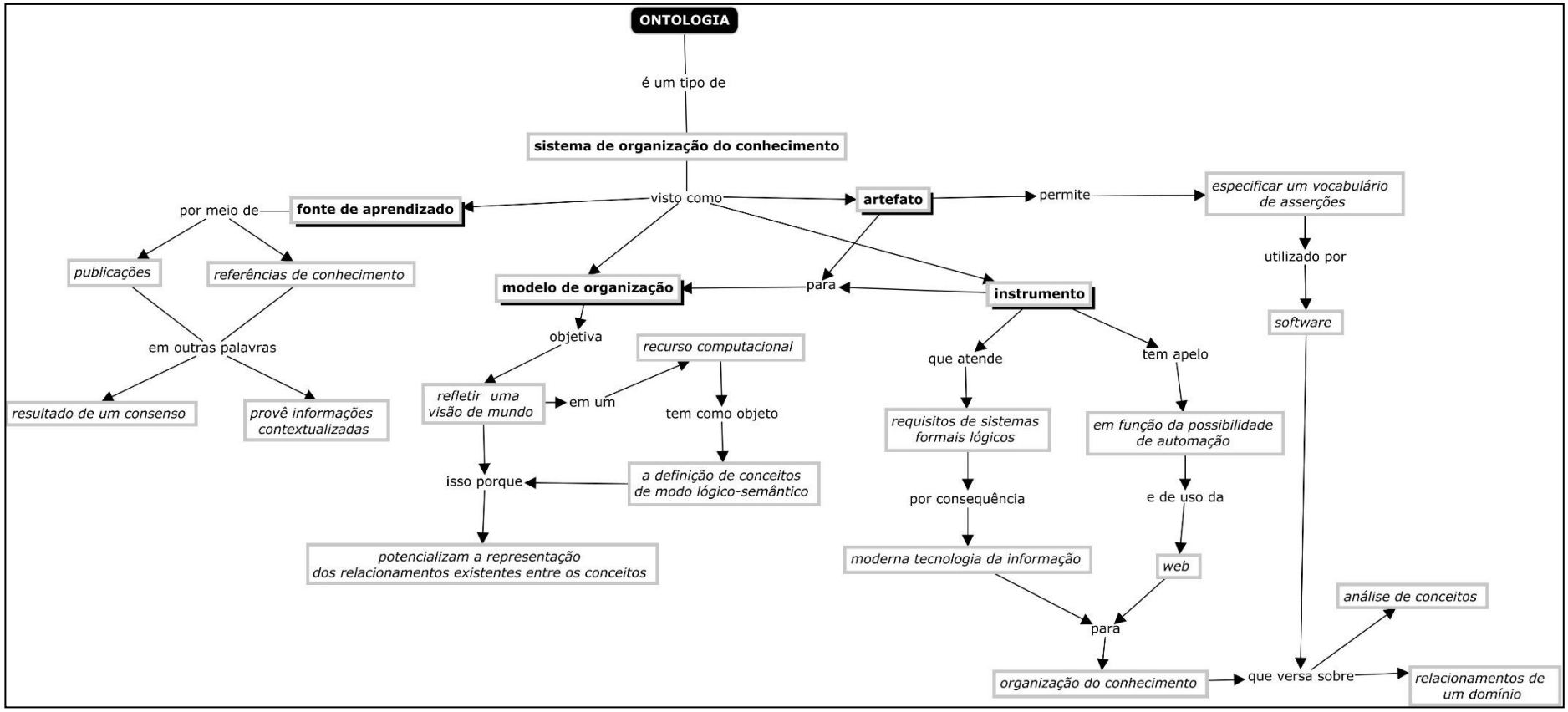
Figura 94 - Concepções sobre ontologia



Fonte: O autor (2022) baseado em dados da pesquisa.

Por meio da leitura da ilustração acima, pode-se compreender que ontologia, para a Ciência da Informação, é um tipo de sistema de organização do conhecimento visto como: **fonte de aprendizado**, por meio de publicações servindo como referência de conhecimento; **modelo de organização**, de algum aspecto da realidade que, por sua vez, reflete uma visão de mundo; e, como **instrumento** de organização da informação e do conhecimento, permite atender aos requisitos de sistemas formais lógicos e da moderna tecnologia da informação; é um **artefato** que objetiva a modelização do conhecimento de algum domínio ou assunto ou, ainda, específico para representações computacionais de algum fenômeno. A extensão dessas concepções conceituais sobre ontologia está apresentada na figura a seguir.

Figura 95 - Concepções estendidas sobre ontologia



Fonte: O autor (2022) baseado em dados da pesquisa.

Durante a exploração da bibliografia selecionada, percebeu-se, demasiadamente, em publicações nacionais e internacionais, referências aos conceitos de ontologia, propostos por Nicola Guarino e Thomas Gruber (CAFÉ; SANTOS; BARROS (2015); CAMPOS; CAMPOS; CAMPOS (2010); DZIEKANIAK (2010); FERNEDA; DIAS (2013); FERNEDA; DIAS (2017); KORNYSHOVA; DENECKÈRE (2010); MENDES NETTO; LIMA (2017); PANSANI JUNIOR; FERNEDA (2016); RAMALHO; ZAFALON; OUCHI (2014); SILVA; RIBEIRO; RODRIGUES (2019); SOUSA *et al*. (2010); SOUZA; ALMEIDA; SOUZA (2017); YAMOKA (2012); ZAHRA *et al*. (2014)).

Nas pesquisas, enquanto Thomas Gruber (1993) conceitua ontologia como uma especificação formal e explícita de conceitualização, Nicola Guarino (1995) soma à definição de Gruber que o estudo desse sistema "[...] pode beneficiar processos de construção de conhecimento produzindo bases de conhecimento de alta qualidade" (SCHIESSL; BRASCHER, 2011, p. 302). Ainda neste sentido,

> Com alto índice de citação no ISI Web of Knowledge, Thomas R. Gruber e Nicola Guarino são autores que têm contribuído para o desenvolvimento do arcabouço teórico e aplicado das ontologias nos sistemas baseados em representações do conhecimento. Gruber é americano, segundo sua página pessoal é um "inovador em tecnologias" que, aplicando conceitos da Inteligência Artificial, Ciências da Cognição e Design, estuda formas de promover a aprendizagem, colaboração e compartilhamento de conhecimento envolvendo máquinas e pessoas. Para tanto, o autor investe desde o início da década de 1990 no estudo, desenvolvimento e aplicação das ontologias no campo da Inteligência Artificial. Gruber é co-fundador do projeto Siri, um agente inteligente para Internet, que atualmente faz parte da Apple, entre outros. Guarino é pesquisador chefe do Laboratório de Ontologia Aplicada que faz parte do Conselho Nacional de Pesquisa da Itália. Desde 1991 se dedica a estudos interdisciplinares entre as Ciências da Computação, Filosofia e Linguística baseando-se na Lógica como paradigma fundador. (CAFÉ; SANTOS; BARROS, 2015, p. 01)

Cabe destacar que, distante das discussões que consideram ontologia como um tipo de linguagem documentária, em relação ao tesauro e taxonomia, autores como Santos e Brascher (2015), Lucas da Silva e Souza (2016), Vilela e Gottschalg (2018), Silva; Ribeiro e Rodrigues (2019), veem ontologia como um tipo de sistema de organização do conhecimento.

Múltiplos estudos versam sobre as diferenças entre ontologia, linguagem documentária, tesauro e taxonomias (ALMEIDA; MENDONÇA; AGANETTE, 2013; AGANETTE; ALMEIDA; TEIXEIRA, 2016; FERNEDA; DIAS, 2013;

VILELA; GOTTSCHALG, 2018; MOREIRA; SANTOS; NETO, 2014; PASANI JUNIOR; FERNEDA, 2016; PIETERSE; KOURIE, 2014). Em razão do objetivo desta tese, recomenda-se que, para um estudo verticalizado dessas diferenças, leitores interessados neste tipo de discussão possam consultar a bibliografia aqui recomendada, como nas referências citadas anteriormente e no artigo intitulado “Sistemas de Organização do Conhecimento: semelhanças e diferenças”, de Nascimento e Pinho (2019), publicado no periódico Pesquisa Brasileira em Ciência da Informação e Biblioteconomia.

Em relação, também, ao tesauro e à taxonomia, o sistema ontologia, para a Ciência da Computação e a Ciência da Informação, já nasce em um contexto tecnológico, em concepções conceituais e processuais. Apesar do encontro conceitual desses sistemas, face à organização e representação da informação e do conhecimento, por suas características em comum de um domínio de interesse, o desenvolvimento e aplicação de ontologias recebe maior atenção quanto ao aprimoramento desse sistema, pela real necessidade de acompanhar o processo de aperfeiçoamento do contexto das tecnologias digitais. Como exemplo, a aplicação desse sistema está cada vez mais reconhecida em domínios: comércio eletrônico; web semântica; sistemas empresariais; integrações de informações em ambientes digitais; modelagem quantitativa de sistemas físicos; processamento de linguagem natural; engenharia de conhecimento e banco de dados (HERRE, 2013).

Nesse sentido, torna-se fundamental compreender que este sistema, para a CI, é capaz de enriquecer os processos de organização e representação da informação e do conhecimento em ambientes digitais. Assim, ontologias são consideradas como artefatos que possibilitam esse tipo de organização e representação computacional, de quaisquer domínios, por meio de regras de inferências definidas durante seu desenvolvimento (MOREIRA; SANTOS NETO, 2014; RAMALHO; ZAFALON; OUCHI, 2014; SANTOS; CORREA; LAPA, 2013).

Rocha, Campos e Costa (2016) discutiram a aplicação de ontologias em razão do tratamento e processamento de dados, especificamente dados abertos. Os autores apontaram a ontologia como um agente facilitador de:

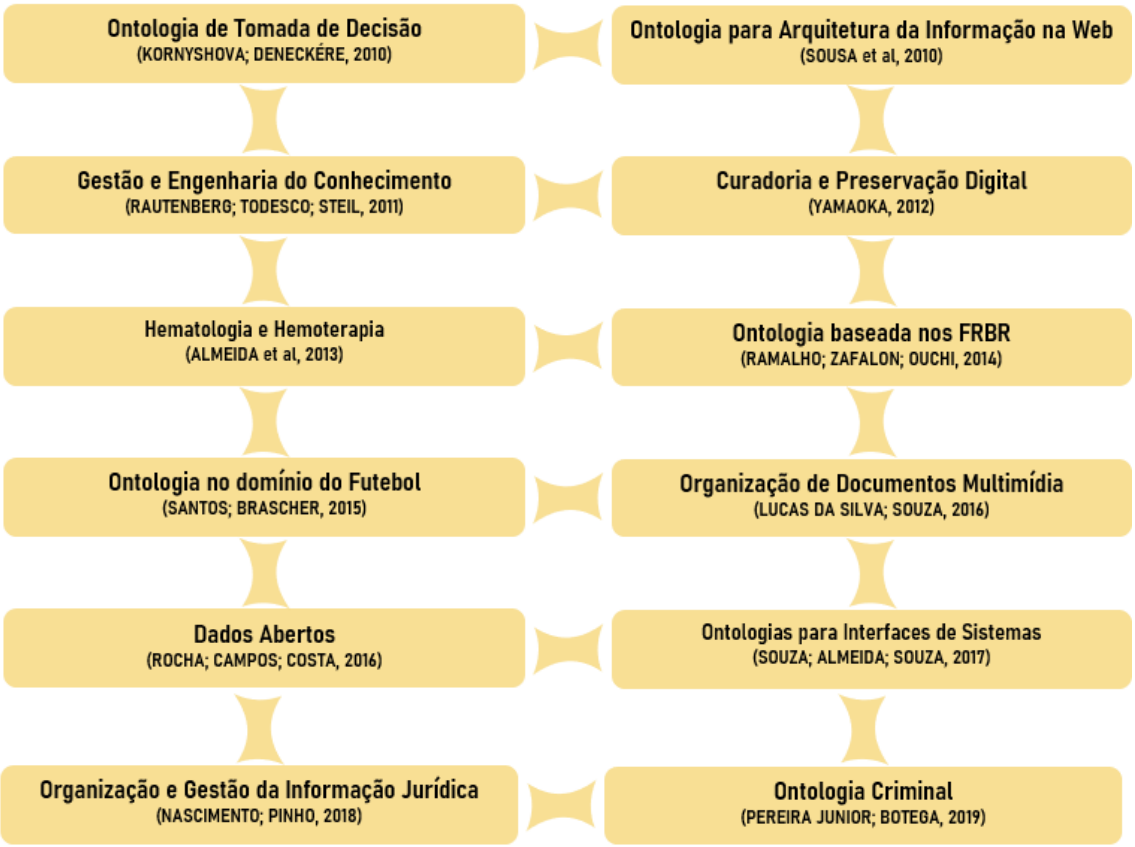
a) Representar dados completos e primários, uma vez que pode utilizar URI's[70] que representam uma determinada instância na web;
b) Representar dados atuais, contanto que seja um sistema em constante atualização;
c) É uma ferramenta importante para tornar uma informação acessível;
d) Ontologia preconiza dados processáveis por máquinas, considerando sua função de produção do conhecimento através de inferências.

Vista como um sistema potencial para recuperação e acesso às informações, dois campos recebem maior destaque quanto desenvolvimento e aplicação de ontologias, são eles: o domínio biomédico, como os estudos de Aganette; Almeida e Teixeira (2016), Mendonça e Almeida (2014; 2015), Silva, e Ribeiro e Rodrigues (2019), e o domínio jurídico, como os estudos de Nascimento, Corrêa e Pinho (2019), Nascimento e Pinho (2018a; 2018b) e Pereira Junior e Botega (2019).

De acordo com os extratos analisados, tal como visto nos sistemas tesauro e taxonomia, a aplicação de ontologias foi percebida em múltiplos contextos. Essas aplicações estão apresentadas na figura a seguir.

[70] *Uniform Resource Identifier* (URI), em português: Identificador Uniforme de Recurso.

Figura 96 - Exemplos de ontologia



Fonte: O autor (2022) baseado em dados da pesquisa.

As características inerentes à construção de ontologias requerem discussões por parte dos profissionais responsáveis pelo seu desenvolvimento, que envolvem desde a ideação das ontologias à etapa de validação, disponibilização, acesso e uso do sistema. O processo metodológico, adotado por esses profissionais, ora apresenta similaridades no percurso trilhado para desenvolvimento do sistema, ora inserem critérios particulares para modelagem da ontologia.

No estudo publicado por Mendonça e Almeida (2014), intitulado "*Princípios metodológicos para desenvolvimento de ontologias: análise das práticas correntes e proposições de melhorias*", os autores reuniram propostas metodológicas apresentando-as de acordo com seus respectivos autores. São exemplos das propostas mapeadas por Mendonça e Almeida:

a) Metodologia TOVE (GRUNINGER; FOX, 1995[71]);
b) Metodologia ENTERPRISE (USCHOLD; KING, 1995[72]);
c) Menthotology (GÓMEZ-PÉREZ; FERNANDEZ-LOPES; VICENTE, 1996[73]);
d) Método Kactus (BERNARAS; LAREGOTTI; CORERA, 1996[74]);
e) Método Sensus (SWARTOUT *et al*., 1996[75]);
f) Método 101 (NOY; MCGUINESS, 2001[76]);
g) Método CYC (REED; LENAT, 2002[77]);
h) Metodologia On-to-knowledge (OTKM) (SURE; STAAB; STUBER, 2003[78]);
i) Metodologia UP for Ontology (UPON) (DE NICOLA; MISSIKOFF; NAVIGLI, 2009[79]);
j) Metodologia NeON (SUARÉZ-FIGUEROA, 2010[80]);
k) Metodologia MFPFO (LIM; LIU; LEE, 2011[81]);

---

[71] GRUNINGER, M.; FOX, M. S. **Methodology for the Design and Evaluation of ontologies**. *In*: Workshop on Basic Ontological Issues in Knowledge Sharing (IJCAI). [*S. l.; s. n*.], 1995.
[72] USCHOLD, M.; KING, M. **Towards a methodology for Building ontologies**. [*S.l.: s.n*.], 1995. Disponível em: https://citeseer.ist.psu.edu/uschold95toward.html. Acesso em: 11 mar. 2014.
[73] GÓMEZ-PÉREZ, A.; FERNÁNDEZ, M.; VICENTE, A. J. **Towards a method to conceptualize domain ontologies**. *In*: Workshop on Ontological Engineering. Budapest: ECAI, 1996.
[74] BERNARAS, A.; LARESGOITI, I.; CORERA, J. **Building and reusing ontologies for electrical network applications**. *In*: THE EUROPEAN CONFERENCE ON ARTIFICIAL INTELLIGENCE. 1996. **Proceedings**... [S.l.]: ECAI, 1996. p. 298-302.
[75] SWARTOUT, B. *et al*. **Toward distributed use of large-scale ontologies**. 1996. Disponível em: https://ksi.cpsc.ucalgary.ca/KAW/KAW96/swartout/Banff_96_final_2.html. Acesso em: 13 de mar. 2014.
[76] NOY, N. F.; MCGUINNESS, D. L. **Ontology Development 101:** a guide to creating your first ontology. [*S.l.: s.n.*], 2001. Disponível em: https://protege.stanford.edu/publications/ontology_development/ontology101.pdf. Acesso em: 06 nov. 2021.
[77] REED, S. L.; LENAT, D. B. **Mapping Ontologies into Cyc**. [*S.l.: s.n.*], 2002. Disponível em: https://www.aaai.org/Papers/Workshops/2002/WS-02-11/WS02-11-010.pdf. Acesso em: 06 nov. 2021.
[78] SURE, Y.; STAAB, S.; STUBER, R. **On-to-knowledge Methodology (OTKM)**. [*S.l.: s.n*.], 2003. Disponível em: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.14.798&rep=rep1&type=pdf. Acesso em: 06 nov. 2021.
[79] DE NICOLA, A.; MISSIKOFF, M.; NAVIGLI, R. A software engineering approach to ontology Building. **Information Systems**, v. 34, p. 258-275, 2009.
[80] SUÁREZ-FIGUEREOA, M. C. NeOn Methodology for Building Ontology Networks: Specification, Scheduling and Reuse. Tese (Doutorado em Inteligência Artificial) – Universidade de madrid, Espanha, 2010.
[81] LIM, S. C. J.; LIU, Y.; LEE, W. B. A methodology for Building a semantically annoted multi-faceted ontology for product Family modelling. **Advanced Engineering Informatics**, v. 25, n. 2, p. 147-161, 2011.

l) Ciclo de Vida Ontológico (SCHIESSL; BRASCHER, 2011).

Neste sentido, para Mendonça e Almeida (2014), na literatura sobre ontologias, “[...] é possível identificar uma série de metodologias e métodos para a construção de ontologias que, em geral, são direcionadas para aplicações em domínios específicos”; essa análise corrobora com a ideia apresentada no parágrafo anterior. Os autores ainda reforçam que, embora

> [...] tais metodologias apresentem algumas semelhanças entre seus aspectos principais e etapas previstas para a construção de ontologias, podem ser percebidas diferenças entre elas, principalmente quando é realizada a comparação entre metodologias direcionadas a domínios de aplicação distintos entre si. (MENDONÇA; ALMEIDA, 2014, p. 433)

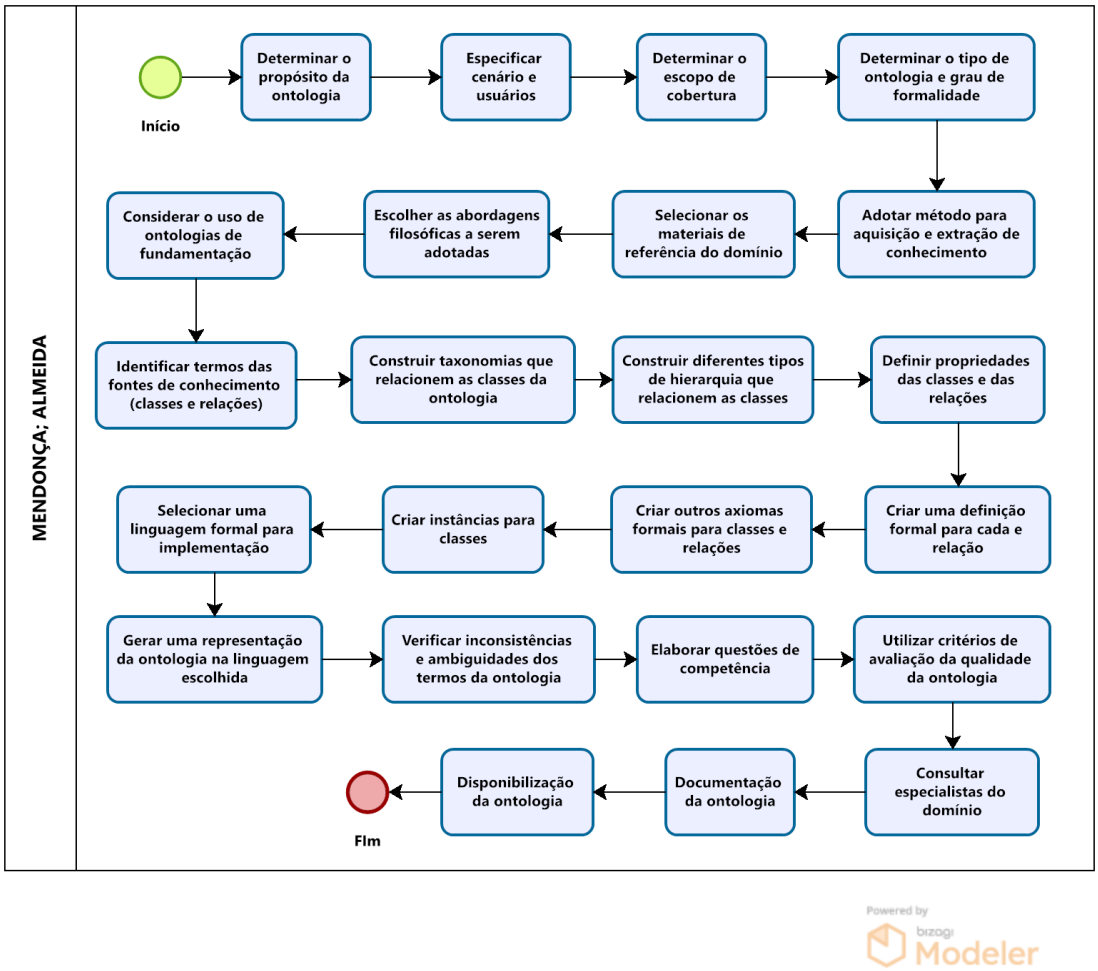
Mendonça e Almeida (2014, p. 436), por meio das percepções extraídas dos métodos e metodologias mapeados e apresentados na lista acima, consideraram uma metodologia própria para construção de ontologias. Para tal, justificaram seu desenvolvimento pelos problemas percebidos nas propostas já existentes. Alguns desses problemas foram pontuados: metodologias e métodos mostram-se superficiais na descrição dos passos para construção de ontologias; algumas abordagens dão mais ênfase em atividades de desenvolvimento; grau de maturidade; muitas das metodologias atuais assumem uma tarefa como ponto de partida na construção da ontologia; as atividades necessárias para reutilização e integração de ontologias são ainda tarefas não triviais e para as quais ainda não existe um padrão amplamente aceito para realizá-las.

Somam-se ainda às análises de Mendonça e Almeida (2014, p. 436) “[...] aspectos ontológicos considerados essenciais no processo de desenvolvimento de ontologias”. Esses processos, segundo os autores, são: delimitação clara e explícita do escopo; uso de ontologias de fundamentação; reuso de termos de outras ontologias; descrição informal e outra formal da ontologia; criteriosidade na transição da descrição informal da ontologia para a descrição formal; elaboração das definições textuais; princípios da mono-hierarquia e da não circularidade; completude mínima necessária do conteúdo ontológico.

Por fim, ao realizarem as análises dos extratos e identificarem as necessidades de melhorias, Mendonça e Almeida (2014) criaram uma

fundamentação própria e mapearam as etapas da metodologia conforme a ilustração a seguir.

Figura 97 - Mapeamento dos processos para construção da ontologia por Mendonça e Almeida



Fonte: O autor (2022) adaptado de Mendonça e Almeida (2014, p. 440).

Para o percurso metodológico, desenhado acima, serão exploradas as tarefas que possuem relação com os objetivos desta tese.

Destarte, os autores classificaram as tarefas de acordo com etapas (MENDONÇA; ALMEIDA, 2014, p. 440) e, assim, as discussões aqui realizadas iniciam-se pela etapa de conceitualização e reutilização, que tem como primeira tarefa *identificar termos das fontes de conhecimento (classes e relações)*; e tem como propósito a descrição da ontologia, em modelo conceitual, no sentido de atender às propriedades das etapas realizadas

anteriormente e dar ênfase ao reuso de termos de ontologias já existentes. Em resumo, as etapas anteriores atendem:

a) Ao planejamento da ontologia, reconhecendo as necessidades; o propósito; os objetivos; o escopo de cobertura; cenários de uso e aplicação; identificação de usuários; e grau de formalidade;
b) À coleta dos termos e o devido comportamento desses no sistema e investigação de métodos para construção de ontologia; análises de textos referentes ao domínio; elaboração de entrevistas com especialistas; seleção dos materiais de referência a serem utilizados;
c) Para fundamentação da ontologia os autores recomendam a adoção de uma ou mais abordagens filosóficas e avaliação do uso da ontologia.

Vista de forma genérica, como os processos anteriores devem ser desenvolvidos, a etapa de conceitualização com reutilização tem o propósito de descrever a ontologia. Nesta etapa, são realizadas as tarefas de reaproveitamento dos termos que, para os autores, “[...] é um dos princípios universalmente aceitos na construção de ontologias” (MENDONÇA; ALMEIDA, 2014, p. 441).

Tão logo, fazem parte dessa etapa de conceitualização com reutilização (MENDONÇA; ALMEIDA, 2014, p. 441):

a) A identificação dos termos do domínio em materiais de referência;
b) A construção de taxonomias relacionando classes do domínio;
c) A construção de outras hierarquias relacionando tais classes;
d) Pesquisar em outras ontologias existentes por um termo semelhante ao termo identificado; para termos não encontrados deve-se reaproveitá-lo do sistema origem;
e) Para desenvolvimento das taxonomias da ontologia deve-se conter classes relacionadas unicamente por meio das relações ontológicas *is_a* e *part_of*;
f) E, as estruturas formais de representação, como: produces; participates_in; has_agent da ontologia.

A etapa seguinte, acerca da metodologia proposta, trata-se da formalização do sistema modelado nos processos realizados anteriormente. Mendonça e Almeida (2014, p. 442) ressaltam, a esse respeito, que “[...] com a evolução dos editores de ontologia, tal como o Protégé 4.3, ela pode ser praticamente toda realizada com uso desses softwares”. A integram as seguintes atividades:

a) Definição formal das propriedades das classes e das relações. Aqui, para completude mínima, os autores sugerem a identificação dos atributos: identificador da classe (*ID*); nome da classe (*label*); identificação da ontologia de origem (*imported from*; *hasOBONamespace*); URI (*hasURI*); sinônimos do termo (*hasSynonym*; *hasExactSynonym*; *hasRelatedSynonym*); termo alternativo (*alternative term*); definição textual (em linguagem natural) (*definition*); elucidação (*elucidation*); exemplo de uso (*example of usage*); definição formal (*SubClass OF*); anotações ou comentários (*comments*);
b) Criação de uma definição formal para cada classe e relação da ontologia que, por meio da definição textual das classes e relações, constrói-se uma definição formal usando relações ontológicas;
c) Construção de outros axiomas formais como forma de assegurar um bom grau de axiomatização do sistema;
d) E, definição de instâncias para cada classe da ontologia.

Após realizarem os processos vistos até aqui, Mendonça e Almeida (2014) apontam uma sexta etapa, concernente à implementação do sistema em uma linguagem formal e computável. Os autores (2014, p. 443) destacam que separaram a etapa de implementação da anterior (formalização) “[...] pelo fato de que a formalização da ontologia deve ser feita de maneira independente e transparente em reação à sua representação em linguagens formais”. Tão logo, como exemplo de atividades inerentes a ela, têm-se:

a) Seleção de uma linguagem formal, geralmente, linguagens lógicas;

b) A geração da representação formal do sistema na linguagem escolhida. Para isso, os autores recomendam o uso de um editor de ontologias que realize essa geração de forma automática.

A validação do conteúdo ontológico é a sétima etapa. As atividades envolvidas são:

a) Verificar e tratar erros de inconsistência, incompletudes e ambiguidades dos termos da ontologia. Pode-se utilizar, por exemplo, classificadores (*reasoners*) disponíveis em editores de ontologia, tal como o Protégé. Esses classificadores são: *FaCT++* e *HermiT*;
b) Especificar questões de competência de validação da ontologia;
c) Uso de critérios de avaliação da qualidade do sistema;
d) Consultar especialistas do domínio para validação do conteúdo ontológico. Aqui tem-se como exemplo a visualização do sistema.

A última etapa da metodologia proposta por Mendonça e Almeida (2014) configura-se como a documentação do sistema, a qual deve incluir os termos a seguir:

a) Elementos de especificação da ontologia. Trata-se de toda ideação realizada nas primeiras etapas, a etapa de *kick-off*;
b) Questões de competência para definição do escopo e validação da ontologia;
c) Relação e breve descrição do sistema que foram reutilizados;
d) Relação dos termos reaproveitados;
e) Identificação dos dicionários consultados;
f) Mapeamento das hierarquias e outras estruturas de representação da ontologia;
g) Os critérios, métodos e técnicas usados na avaliação e validação do sistema;
h) Relação e descrição dos materiais de referência.

A riqueza metodológica percebida na proposta de Almeida e Mendonça (2014) não foi percebida nos extratos consultados e analisados. A literatura se

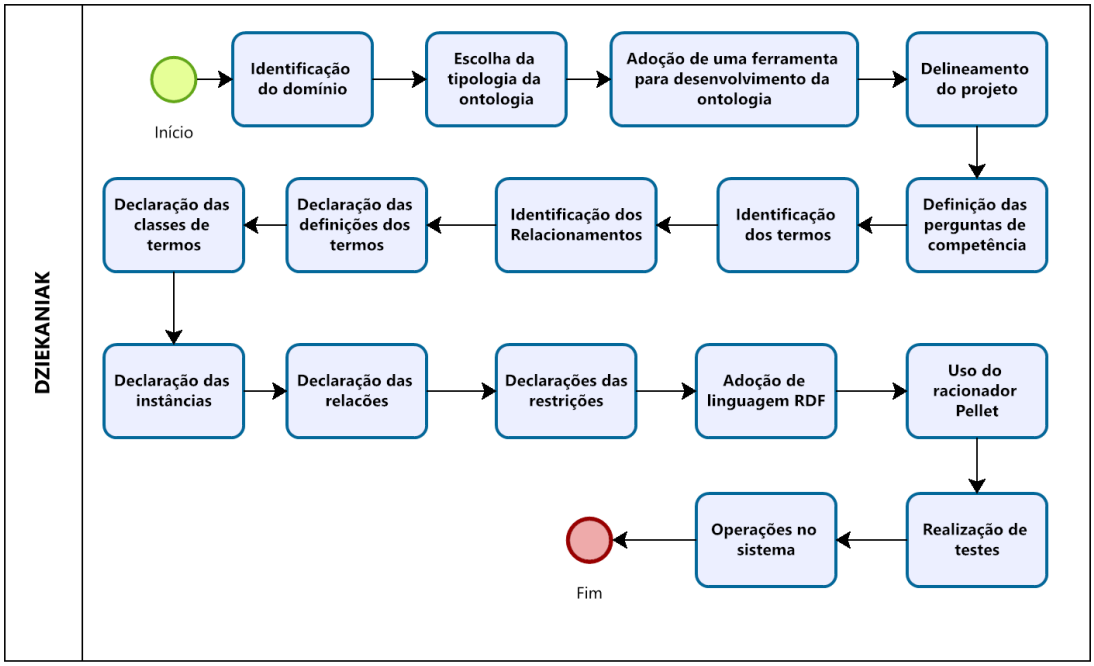
aporta de múltiplas pesquisas publicadas, de métodos já orientados à aplicação e metodologias para construção, acesso e uso. Todavia, mas, de forma tímida, ou quase escassa, não se aprofunda nas operacionalidades realizadas para desenvolvimento, aplicação e uso de sistemas de organização do conhecimento, tais como ontologias.

Contudo, percebe-se que as propostas de percursos metodológicos vistas nos últimos anos ganharam maior destaque quanto à aplicação de processos próprios integrados a outros já existentes. Tais como as metodologias que, de certa forma, consagraram-se no meio computacional e, por isso, cristalizaram-se, também, na Ciência da Informação.

Como exemplo de metodologias demasiadamente utilizadas em trabalhos técnico-científicos, pode-se citar: *Ontology Development 101* (NOY; MCGUINESS, 2001); *On-to-knowledge* (SURE; STAAB; STUBER, 2003); *Methontology* (GÓMEZ-PÉREZ; FERNANDEZ-LOPES; VICENTE, 1996); e a metodologia *Enterprise* (USCHOLD; KING, 1995). Os processos mapeados abaixo integram uma ou mais dessas metodologias no constructo de ontologias.

Dziekaniak (2010) publicou o artigo "*Desenvolvimento de ontologia sobre componentes de ontologias*", onde apresentou (2010, p. 173) "[...] os principais passos no desenvolvimento de ontologia de domínio sobre Componentes de Ontologias". Para tanto, a autora se aportou das metodologias *Methontology*, *101* e *On-to-knowledge.* Os processos apresentados pela autora estão mapeados na ilustração abaixo:

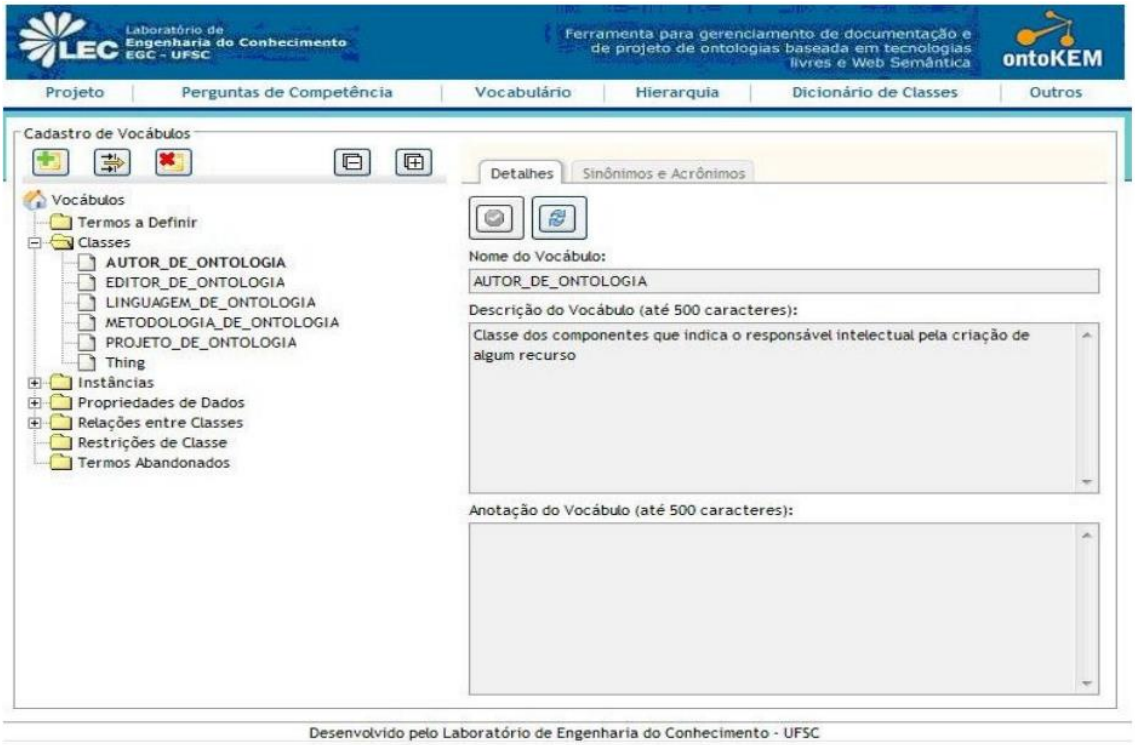
Figura 98 - Mapeamento dos processos para construção da ontologia por Dziekaniak



Fonte: O autor (2022) com base em Dziekaniak (2010).

Após os primeiros processos (identificação do domínio e escolha da tipologia da ontologia), a autora partiu para a seleção de uma ferramenta orientada à edição da ontologia. Uma das ferramentas selecionadas apara o desenvolvimento da ontologia foi o Ontokem, que "[...] é um sistema desenvolvido pelo Laboratório de Engenharia do Conhecimento da Universidade Federal de Santa Catarina (LEC/ UFSC)" (DZIEKANIAK, 2010, p. 176). A estrutura do Ontokem contempla 6 seções, que representam as etapas básicas para construção de ontologias (DZIEKANIAK, 2010). A figura abaixo ilustra as seções do Ontokem.

Figura 99 - Seções do Ontokem



Fonte: Dziekaniak (2010, p. 178).

Além do uso do Ontokem, Dziekaniak (2010, p. 181) utilizou o software Protégé. Neste software, “[...] foram trabalhados os domínios e os ranges e também criadas as restrições do projeto”. Em comparação com a proposta de Mendonça e Almeida (2014), as restrições no processo de Dziekaniak

> [...] servem para limitar e auxiliar no processo de uma pergunta à base ontológica. Neste projeto de ontologia utilizou-se apenas a restrição *some* para “forçar” que todo projeto de ontologia citado nesta ontologia, toda metodologia de ontologia, toda linguagem de ontologia, todo editor de ontologia tenha pelos um autor. Uma vez que para o projeto desenvolvido, a questão autoral foi considerada essencial.

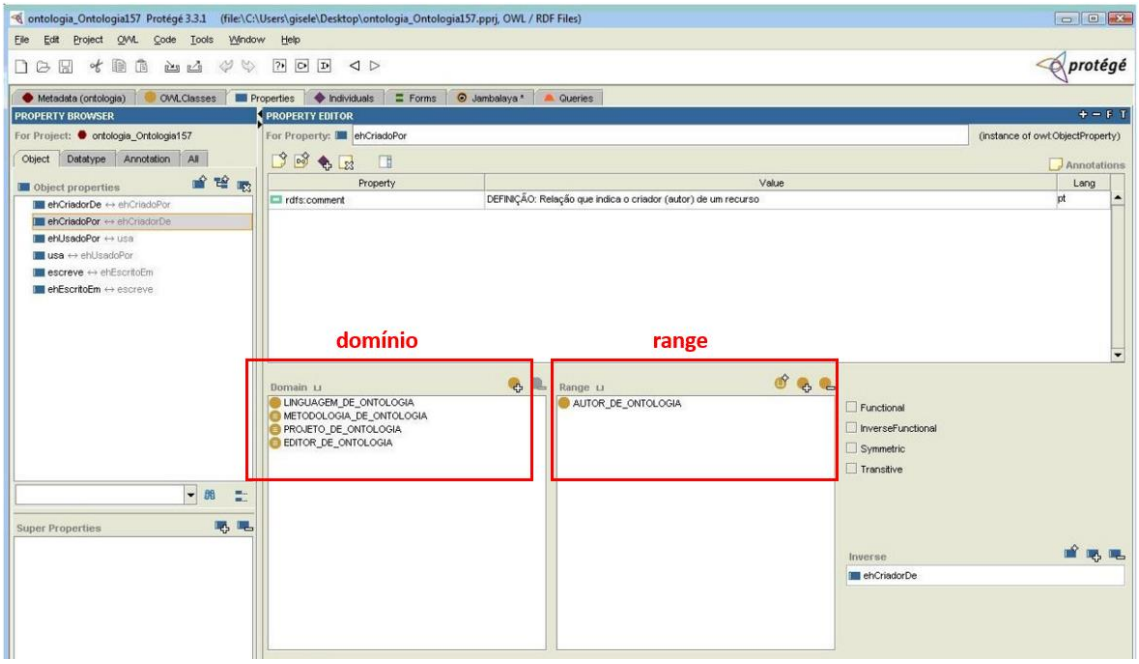
Em sequência aos processos, a atribuição dos domínios e ranges, foi considerada uma outra etapa importante no desenvolvimento da ontologia, vista na seção de tesauros, que se trata dos conceitos herdados da linguagem RDF. Conforme Dziekaniak (2010, p. 181):

> Esta linguagem é adotada nos projetos de ontologias devido à possibilidade de representação entre um relacionamento e os objetos que se relacionam através da tripla sujeito-propriedade-objeto, as quais contém afirmações sobre dada informação. Esta representação

em RDF é reconhecida pela linguagem XML, o que possibilita que a ontologia seja implementada em ambiente web.

A figura abaixo ilustra o exemplo de domínio e range.

Figura 100 - Domínio e range em ontologia



Fonte: O autor (2022) adaptado de Dziekaniak (2010, p. 182).

Para além do Ontokem e *Protégé*, Dziekaniak utilizou um raciocinador denominado *Pellet*. Este, por sua vez, funciona como um *pluggin* que "[...] dialoga com o *Protégé* para realizar inferências na base de conhecimento desenvolvida no projeto" (DZIEKANIAK, 2010, p. 183). Por fim, procedeu-se à etapa de testes com o objetivo de comprovar que a ontologia era capaz de responder às perguntas de competência inferidas no *kick-off* do projeto (DZIEKANIAK, 2010)

Tal como Dziekaniak (2010), Sousa *et al*. (2010), também se aportou da metodologia 101. Essa metodologia é composta por sete etapas não linear, a saber (NOY; MCGUINESS, 2001; SOUSA *et al*., 2010):

I) Determinar o domínio e o escopo do sistema;

II) Considerar o reuso de ontologias já desenvolvidas;

III) Ordenar os termos importantes;

IV) Definir as classes e suas respectivas hierarquias;

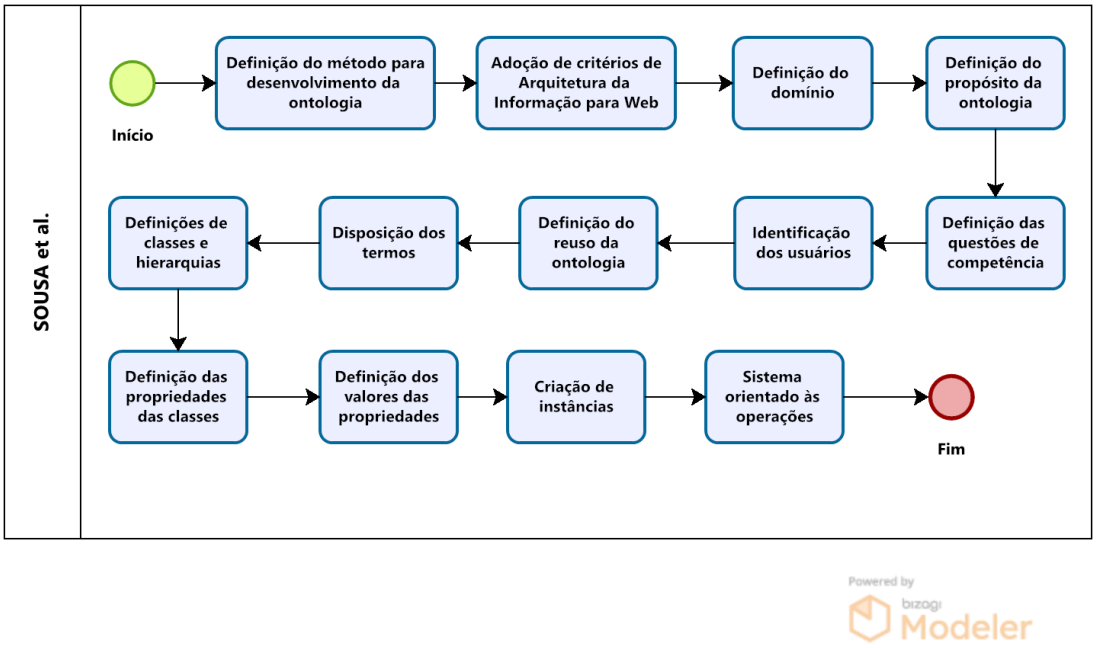
V) Definir as propriedades das classes;

VI) Definir os valores das propriedades das classes;

VII) Criar as instâncias.

Contudo, Sousa *et al*. (2010) adaptaram as etapas vistas acima e criaram um percurso metodológico próprio. Essa adaptação deve-se, principalmente, ao objetivo do desenvolvimento da ontologia, considerando os principais conceitos do domínio da Arquitetura da Informação para Web. Assim, a figura abaixo representa o percurso metodológico trilhado pelos autores.

Figura 101 - Mapeamento dos processos para construção da ontologia por Sousa et al



Fonte: O autor (2022) com base em Sousa *et al*. (2010).

Os dois primeiros processos fazem parte do início do planejamento da construção da ontologia. O terceiro, da definição do domínio. Partiu-se de dois conceitos fundamentais para Arquitetura da Informação para Web, sendo esses o sistema (organização; navegação; rotulação; e busca) e a metodologia que buscam a organização das fases para o desenvolvimento da arquitetura para web. O propósito da ontologia e as questões de competência dialogam com os objetivos para os quais o sistema foi construído. Como exemplo das questões de competências, têm-se: “[...] quais os componentes da AI que devem compor os sites? Qual a função dos sistemas de organização, navegação, rotulação e

busca nos sites?" (SOUSA *et al*., 2010, p. 10). Consideram-se, ainda, os usuários do sistema, isto é, todos aqueles que tenham interesse sobre o assunto. Os autores citam como exemplo os professores.

O reuso de ontologia é outro tema central nos percursos metodológicos consultados para esta tese. Neste sentido, Sousa *et al*. (2010) conceberam o sistema de forma que a modelagem propriamente dita possa ser reutilizada à criação de novas ontologias relacionadas à Arquitetura da Informação para Web. Ademais, pode ainda ser reutilizada com o uso de um *framework*, como exemplo a linguagem OWL, a qual também os autores se disponibilizaram a compartilhar mediante contato.

Enumerar os termos da ontologia e definir classes e suas respectivas hierarquias são outras etapas do percurso metodológico. A definição de classes foi efetuada no *Protégé*, onde todas elas herdam de uma classe nativa (*thing*). A definição das propriedades de cada uma foi efetuada em dois tipos: *object properties* ou *datatype properties*.

Na sequência, chega-se ao processo de definir os valores das propriedades. Para Sousa *et al*. (2010, p. 15): "[...] cada propriedade deve ter uma descrição diferente como o tipo de valor (funcional, inverso, simétrico ou transitivo), a cardinalidade, e outros recursos que o campo pode ter."

A criação das instâncias antecede o último processo do percurso metodológico. Aqui, são criadas as instâncias individuais para cada classe, "[...] para que por meio dos relacionamentos entre essas classes possam ser definidas as consultas que retornam resultados como respostas a questões de competência" (SOUSA *et al*., 2010, p. 16).

Exemplos de definições: "Componentes; Estruturas_de_Representação; Sistemas_de_Busca; Sistema_de_Organização; Sistemas_de_Rotulação [...]" (SOUSA *et al*., 2010, p. 16).

Como contribuições da modelagem da ontologia e seu compartilhamento os autores destacaram:

> - Servir como base para o conhecimento do domínio da Arquitetura da Informação para Web, possibilitando o reuso e compartilhamento dos conceitos e seus relacionamentos. A Arquitetura da Informação para Web é um tema ainda novo e pouco conhecido, com pouca literatura disponível, sobretudo em língua portuguesa. Assim, a InfoArch também servirá de fonte de consultas podendo inclusive ilustrar de forma gráfica seus conceitos e

> relacionamentos, o que vem a facilitar o seu entendimento, principalmente em estudos feitos por iniciantes no assunto;
> - Ajudar no desenvolvimento de sites que contemplem os componentes da Arquitetura da Informação para Web, possibilitando aos usuários destes, terem mais facilidade para encontrar as informações que desejam. Neste caso, a ontologia também representa mais uma fonte de consulta, podendo ser utilizada pela equipe de desenvolvimento para sanar dúvidas sobre como utilizar os componentes da Arquitetura da Informação para Web, o que até então, só seria capaz por meio da literatura disponível e dos arquitetos da informação;
> - Auxiliar no ensino da disciplina de Arquitetura da Informação para Web, que já faz parte de Programas de Pós-Graduação relacionados a áreas como Ciência da Informação e Design Gráfico. Como também, em cursos ministrados por professores pesquisadores do assunto e/ou por empresas de desenvolvimento de sites, tanto para treinar seus desenvolvedores quanto para agregar os componentes (sistemas) da Arquitetura da Informação para Web em seu processo de desenvolvimento. (SOUSA *et al*., 2010, p. 19).

Um destaque para a ontologia projetada e construída por Sousa *et al*. (2010), de acordo com as contribuições apresentadas, é o despertar para a socialização do conhecimento aplicado tanto para os profissionais relacionados à Arquitetura da Informação para Web "[...] na área acadêmica em pesquisa, ensino ou extensão, quanto na área coorporativa com as pessoas e empresas responsáveis pelo desenvolvimento de *sites*" (SOUSA *et al*., 2010, p. 20, grifo do autor).

Como visto até aqui, ontologias podem ser desenvolvidas de acordo com o escopo e planejamento realizado pela equipe responsável por sua construção. Tem-se como exemplo a pesquisa de Mendonça e Almeida (2014), que adotou métodos próprios, por considerar que os já existentes (ou consolidados) apresentam suas necessidades de melhorias na cadeia de processos. E, por outro lado, existem pesquisas que corroboram com os percursos metodológicos já desenvolvidos, e o tomam como guia para desenvolvimento de ontologias; ou pesquisas que se aportam de processos de metodologias existentes, para aprimorar um percurso metodológico próprio.

Ainda como exemplo de pesquisas que idealizam ontologias, por meio da integração de processos de múltiplas metodologias, pode-se citar Rautenberg, Todesco e Steil (2011), com a publicação da pesquisa "Uma ontologia para instrumentos da Gestão do Conhecimento e Agentes da Engenharia do Conhecimento". Ao método utilizado no desenvolvimento da

ontologia proposta pelos autores, somam-se as combinações de artefatos metodológicos oriundos da (RAUTENBERG; TODESCO; STEIL, 2011, p. 116):

a) *On-to-Knowledge* (SURE; STAAB; STUDER, 2003): contribui para a especificação dos requisitos da ontologia, por meio do emprego de questões de competência como modo simples e direto para confirmar o propósito e o escopo de uma ontologia. Tal fato permite identificar, antecipadamente, conceitos, propriedades, relações e instâncias de uma ontologia;
b) *Methontology* (GÓMEZ-PÉREZ; CORCHO; FERNÁNDEZ-LÓPEZ, 2004[82]): por meio de uma rica gama de artefatos, contribui na documentação e na verificação de ontologias.
c) *Ontology Development 101* (NOY; MCGUINNESS, 2001): contribui com uma visão clara de como se dá um processo iterativo para o desenvolvimento de ontologias.

Os autores, portanto, apresentaram um percurso metodológico orientado por quatro fases, a saber (RAUTENBERG; TODESCO; STEIL, 2011, p. 118):

Fase I – Especificação: atividades cujo propósito é discernir sobre os custos do desenvolvimento de uma ontologia. É inerente a esta fase: identificar o escopo e o propósito do sistema; as fontes de conhecimento disponíveis para consultas e uso; e considerar o reuso de termos e ontologias existentes;

Fase II – Aquisição do Conhecimento: atividades de maior interação com os especialistas de domínio para abstrair os elementos de uma ontologia. Esta demanda maior esforço aplicado. Faz parte de seus processos: gerar questões de competência; agregar elementos reutilizáveis; listar termos; definir as classes; as relações; as propriedades; as restrições; e, por fim, criar as instâncias;

Fase III – **Implementação**: atividades de menor interação com os especialistas de domínio, reservada às tarefas de cunho técnico/ tecnológico. Quando se explanou que o desenvolvimento de SOC depende da comunhão de múltiplos especialistas para garantir a qualidade do tratamento, organização

[82] GÓMEZ-PÉREZ, A.; CORCHO, O.; FENÁNDEZ-LÓPEZ, M. **Ontological Engineering:** with examples from the áreas of knowledge management, e-commerce and the Semantic web. Heidelberg: Springer, 2004.

e recuperação da informação, tem-se esta fase como um dos exemplos em que profissionais da Computação são imprescindíveis para elaboração e execução técnica desta fase da construção de ontologia. São exemplos de processos que constituem esta fase: valorar as propriedades; as relações; as restrições;

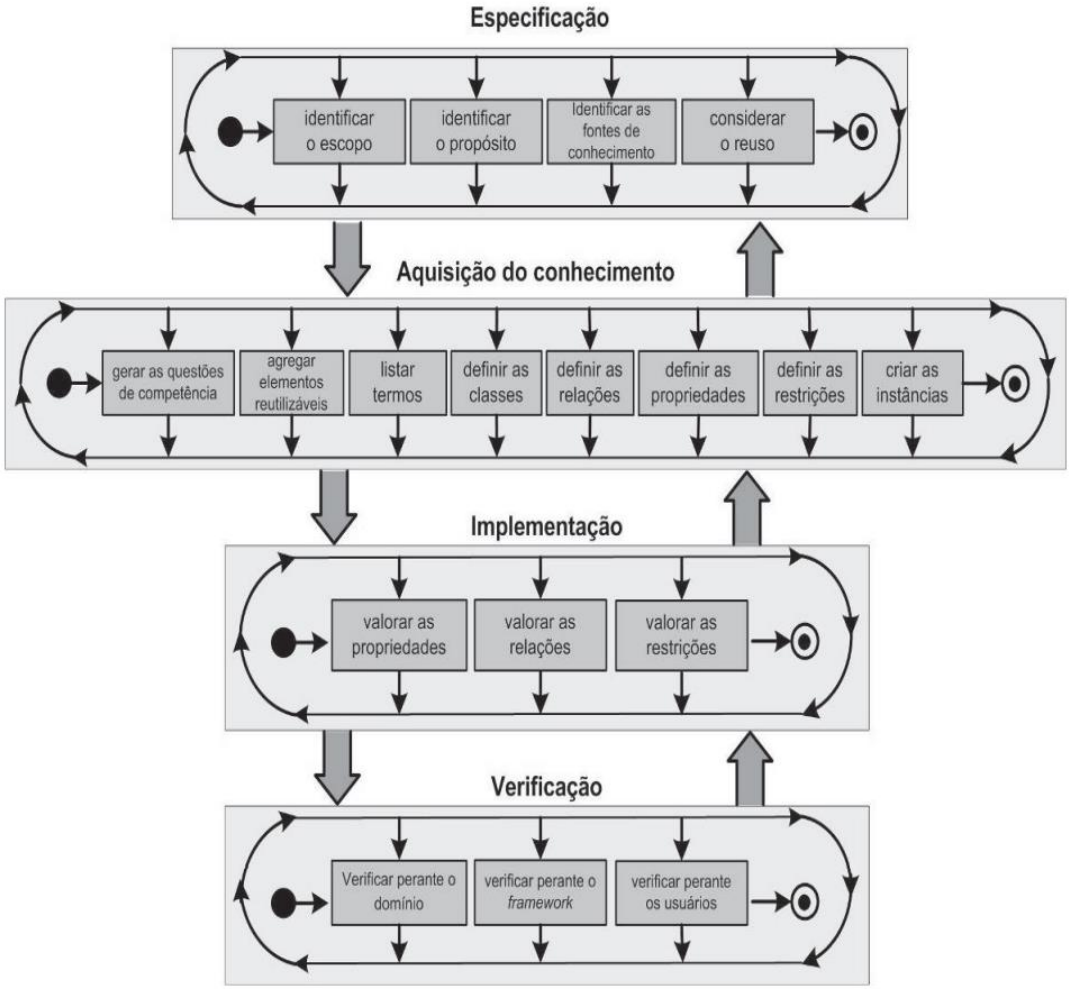
Fase IV – Verificação: atividade que resgata a interação com especialistas de domínio para averiguar se a ontologia desenvolvida atende aos requisitos especificados. Esta é a fase de validação do funcionamento e uso da ontologia. Seus processos são: verificar perante o domínio; o framework; e os usuários.

A Fase de Verificação é um dos destaques principais do percurso de Rautenberg, Todesco e Steil (2011). Ao considerar a metodologia empregada, os autores ressaltam que se deve verificar a ontologia pelos três processos indicados anteriormente e que por eles deve-se verificar o sistema:

> 1. perante o domínio da aplicação de uma ontologia prevenir-se das inconsistências em relação ao entendimento aceito expresso nas fontes de conhecimento coletadas;
> 2. frente ao framework de referência (o propósito, o escopo e as questões de competência da ontologia) é preciso avaliar a consistência da ontologia de acordo com os requisitos levantados; e
> 3. considerando os possíveis usuários, deve-se questionar os indivíduos quanto à utilidade, à precisão e à cobertura da ontologia na explicitação do conhecimento representado.

A ilustração a seguir apresenta a elucidação do mapeamento das fases e processos relacionados dispostos acima.

Figura 102 - Mapeamento dos processos para construção da ontologia por Rautenberg, Todesco e Steil

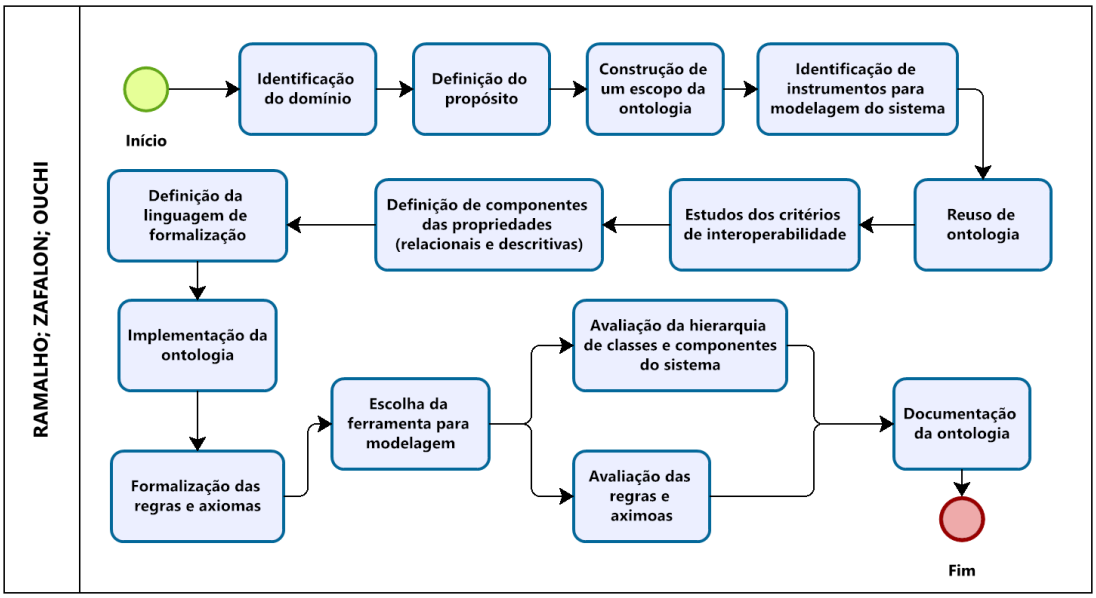


Fonte: Rautenberg, Todesco e Steil (2011).

A Fase de Verificação do método de Rautenberg, Todesco e Steil (2011) reflete a importância de considerar o usuário como elemento estratégico para desenvolvimento de ontologias. Essa consideração pode ser vista também na pesquisa de Ramalho, Zafalon e Ouchi (2014).

A pesquisa intitulada “Ontologia baseada nos FRBR: proposta de aplicação em catálogos online” teve como objetivo o desenvolvimento de um protótipo de ontologia, baseada nas relações de entidades dos Requisitos Funcionais dos Registros Bibliográficos (RAMALHO; ZAFALON; OUCHI, 2014, p. 4134). Para tanto, os autores apresentaram o percurso metodológico desenhado a seguir.

Figura 103 - Mapeamento dos processos para construção da ontologia por Ramalho, Zafalon e Ouchi



Fonte: O autor (2022) com base em Ramalho, Zafalon e Ouchi (2014).

O percurso desenhado acima foi extraído da pesquisa de Ramalho, Zafalon e Ouchi (2014), os quais se basearam no método proposto por Ramalho (2010). Assim, após as etapas de identificação do domínio e definição do propósito da ontologia, realizou-se um estudo com o objetivo de identificar instrumentos utilizados pela Ciência da Informação como substrato teórico. Assim, “[...] as entidades e relacionamentos previstos nos FRBR serviram de arcabouço teórico para a definição dos componentes estruturais da ontologia” (RAMALHO; ZAFALON; OUCHI, 2014, p. 4140).

Na fase de reuso de termos e componentes utilizados para desenvolvimento da ontologia, os autores identificaram vocabulários e outras

ontologias que foram desenvolvidas com a mesma temática da proposta da pesquisa. “Porém, devido ao escopo baseado nas entidades e relacionamentos previstos nos FRBR, recorreu-se ao reaproveitamento de padrões de metadados já sedimentados na área de Catalogação” (RAMALHO; ZAFALON; OUCHI, 2014, p. 4141).

A nomeação dos componentes para formalização da ontologia também foi etapa consideravelmente estratégica no desenvolvimento do protótipo da ontologia:

> Em relação à nomeação dos componentes da ontologia, com o objetivo de facilitar a compreensão e evitar possíveis problemas de interoperabilidade, seguiu-se a convenção de não utilizar acentuação, espaços e caracteres especiais. Definiu-se, também, a utilização de letras maiúsculas no início e em todas as palavras subsequente contidas no nome das classes e subclasses, notação comumente denominada como *CamelBack*. O mesmo procedimento foi adotado para representar as propriedades relacionais da ontologia, com exceção do uso de letras maiúsculas no início de seus nomes, diferenciando-as das classes. Quanto às propriedades descritivas, como o intuito de facilitar a assimilação dos componentes da ontologia, utilizou-se como padrão a denominação prevista nos FRBR, em língua inglesa, sem espaçamento. (RAMALHO; ZAFALON; OUCHI, 2014, p. 4141)

Quanto à aplicação técnica para formalização do sistema, os autores adotaram linguagens processáveis por máquinas. Como a “OWL-DL/XML”, que foi utilizada “[...] para implementação, dada a aderência aos padrões do W3C e às características que possibilitam maior decidibilidade lógica/computacional.” RAMALHO; ZAFALON; OUCHI, 2014, p. 4141). Na condição de ferramenta de apoio para desenvolver a ontologia e, assim, formalizar as regras e axiomas, os autores utilizaram o software *Protégé* mais o motor de inferência padrão da ferramenta FaCT++ (RAMALHO; ZAFALON; OUCHI, 2014).

Após a escolha da ferramenta de modelagem da ontologia, a pesquisa seguiu-se para duas etapas que se auto complementam. Como visto no desenho acima, as etapas são aquelas voltadas à avaliação do protótipo. Para os autores (2014, p. 4142):

> - Avaliação da hierarquia de classes e componentes da ontologia: realizada a partir dos critérios da metodologia OntoClean, (re)adequando ontologicamente os componentes estruturais da ontologia até a obtenção do modelo desejado;
> - Avaliação das regras e axiomas: efetuada por meio do julgamento do resultado de buscas que se aproximem das necessidades informacionais de usuários em potencial, simuladas

> para avaliar a capacidade representação, recuperação e realização de inferências automáticas de acordo os objetivos propostos.

Por fim, de acordo com o percurso apresentado, a atividade final do desenvolvimento da ontologia foi a de documentação. Para isso, utilizaram-se "[...] os próprios recursos da ferramenta Protégé (a partir da inserção de anotações em seus componentes)" (RAMALHO; ZAFALON; OUCHI, 2014, p. 4141).

A pesquisa "Ontoforinfoscience: uma metodologia detalhada para construção de ontologias e sua aplicação no domínio da biomedicina", de autoria de Mendonça e Almeida (2015, p. 3), apresentou uma metodologia para construção de ontologia, a qual pode auxiliar especialistas em Organização do Conhecimento "[...] a superar problemas relativos ao jargão técnico e a questões lógicas e filosóficas que envolvem o desenvolvimento de ontologias". Tão logo "[...] a OntoForInfoScience representa uma iniciativa em direção a um maior entendimento de termos e detalhes técnicos (lógicos e filosóficos) do processo de desenvolvimento de ontologias por parte dos cientistas da informação" (MENDONÇA; ALMEIDA, 2015, p. 3).

Apesar da diferença quanto à temática e ao objetivo da pesquisa de Ramalho, Zafalon e Ouchi (2014), o percurso apresentado por Mendonça e Almeida (2015) assemelha-se ao dos autores anteriormente citados. Em geral:

> Na prática, a metodologia OntoForInfoScience foi utilizada e testada na construção de uma ontologia no domínio da hematologia e hemoterapia – a HEMONTO – e, nesse processo, mostrou-se eficiente quanto ao propósito de gerar uma representação formal adequada. Tal metodologia utiliza uma linguagem simples, que permite a compreensão das atividades executadas ao longo do processo de construção, sem perder a capacidade de expressividade do uso de linguagens lógicas. Algumas funcionalidades da OntoForInfoScience que são úteis e contribuem para o desenvolvimento de ontologias são: i) explicação sobre o significado e o preenchimento das propriedades descritivas e lógicas das classes e relações; ii) incentivo o uso de ontologias de fundamentação como ponto de partida do desenvolvimento ontológico; iii) detalhamento do processo de criação de definições textuais e formais das classes de uma ontologia; iv) explicação sobre o significado de uma relação ontológica e de que forma é possível caracterizar os tipos de relação ontológica para usá-los adequadamente na representação do domínio; v) incentivo a colaboração como um aspecto essencial na etapa de conceitualização, com o uso de ferramentas colaborativas; vi) explicação de formas para pesquisa por termos relacionados ao domínio para importação na ontologia [...] (MENDONÇA; ALMEIDA, 2015, p. 15).

Pode-se observar, nos mapeamentos e modelagens dos processos sobre desenvolvimento de ontologias, a intensidade dos esforços manuais para construção deste sistema, isso vale desde a etapa do planejamento; levantamento e tratamentos de dados; o processamento dos dados nas ferramentas de apoio; o uso de ferramentas especializadas para organização e disposição das informações; de aparatos técnicos que facilitem a estrutura da ontologia; a aplicação e uso do sistema etc.

Apesar do sistema ontologia nascer em berço tecnológico, tal como taxonomia, ser rico em propriedades técnicas e dotado de procedimentos que carecem do conhecimento computacional, nos últimos anos, pesquisas vêm sendo desenvolvidas no sentido de minimizar os esforços manuais aplicados para construção, aplicação e uso de ontologias, principalmente no que tange as etapas de levantamento e uso de conhecimento publicado, isto é, a consulta e uso de textos.

Nascimento, Corrêa e Pinho (2019, p. 137) corroboram com a ideia acima, quando enfatizam, também, que:

> Na tentativa de minimizar parte desses esforços surge o processo de aprendizagem de ontologias a partir de textos, a qual compreende as fases de extração dos termos relevantes, a descoberta dos sinônimos entre esses termos, a ordenação hierárquica dos termos, ou seja, a geração de taxonomias, e por fim, estabelecer os relacionamentos existentes entre eles.

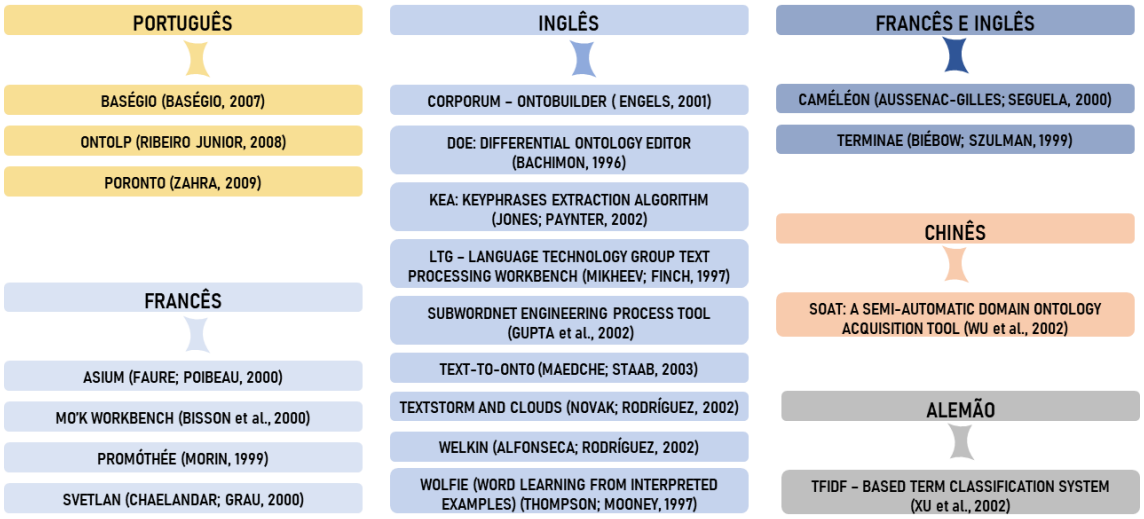
Estudos como os de Ferneda e Dias (2017), Nascimento, Corrêa e Pinho (2019), Pansani Junior e Ferneda (2016), Rocha, Campos e Costa (2016) e Zahra *et al*. (2014) reforçam a ideia do uso de mecanismos automáticos para desenvolvimento de ontologias ou o uso de ontologias para processamento automático de atividades como a de indexação. Assim, os autores somam, à discussão do desenvolvimento de ontologias, as práticas tecnológicas de reuso; interoperabilidade; mineração de textos; formalização das estruturas taxonômicas geradas automaticamente; e do uso de ontologia para processamento da indexação automática etc.

Nesse sentido, a pesquisa de Zahra *et al*. (2014, p. 6) discutiu sobre métodos de aprendizagem de ontologias. Para os autores,

> [...] na aquisição de conhecimentos, os engenheiros/especialistas possuem dificuldades em capturar e representar o conhecimento, mais precisamente, elicitar conhecimentos a serem representados na ontologia e, finalmente, formalizar e implementar este conhecimento em uma linguagem interpretável por máquina. Deste modo, devido ao tamanho, complexidade e dinamicidade de um determinado domínio, métodos de aprendizagem de ontologias têm sido desenvolvidos com o objetivo de reduzir esforços na aquisição de conhecimentos. (...) Para dar suporte ao processo de aquisição do conhecimento e para a construção semiautomática de ontologias a partir de textos, várias ferramentas foram desenvolvidas, cada uma utilizando técnicas e métodos diferentes. Devido à importância do uso destas ferramentas, faz-se necessário conhecê-las, entender quais componentes da ontologia que podem ser construídos em cada ferramenta, assim como quais estão disponíveis para uso e a sua facilidade.

Assim, Zahra *et al*. (2014) apresentaram ferramentas para aprendizagem de ontologias para língua portuguesa e línguas estrangeiras. Os autores se basearam em Gómez-Pérez e Manzanzo-Macho (2003)[83] para levantamento das ferramentas em línguas estrangeiras. O esquema abaixo reúne as ferramentas encontradas pelos autores:

Figura 104 - Ferramentas para aprendizagem de ontologias



Fonte: O autor (2022) com base em Zahra *et al*., 2014.

Zahra *et al*. (2014, p. 18) ressaltaram que "[...] a aprendizagem de ontologia, a partir de textos, pode ser útil para a construção de uma ontologia

[83] GÓMEZ-PÉREZ, A.; MANZANZO-MACHO, D. **A survey of ontology learning methods and techniques**. Espanha: Universidad Politécnica de Madrid, 2003.

ou para o enriquecimento de uma já existente, para diferentes finalidades". Ainda para os autores

> Há várias ferramentas para aprendizagem de ontologias a partir de textos. Muitas delas resultam de trabalhos acadêmicos e poucas delas estão disponíveis para download e continuam a receberem incrementos. Além disso, algumas ferramentas não são de fácil utilização, principalmente quando empregadas por pessoas que não são da área da computação. Tais ferramentas requerem o entendimento de algoritmos, métricas e fórmulas específicas e são destinadas principalmente a realização de experimentos científicos e, em última análise, a engenheiros do conhecimento (mas não a usuário que são especialistas do domínio). Outro entrave é que as ferramentas disponíveis representam apenas algumas etapas de todo o processo e desenvolvimento de uma ontologia, não estando inseridas em uma metodologia maior, destinada ao desenvolvimento de ontologias. Zahra *et al*. (2014, p. 18)

A pesquisa apresentada por Nascimento, Corrêa e Pinho (2019), que teve como objetivo propor e discutir um percurso metodológico para a construção de ontologias do domínio jurídico, trouxe um caminho em que se adotaram técnicas de mineração de textos, mais especificamente de aprendizado de ontologias. Para tanto, os autores tomaram por base os processos de mineração de texto de Gazje (2010) e discutiram as seguintes etapas:

a) **Transformação do texto**: consiste na transformação do documento para um formato de texto;
b) **Separação de palavras**: é o tratamento para identificar as características importantes do texto;
c) **Stemming**: tratamento dos termos por meio da remoção de sufixos e a identificação nos formatos mais puros;
d) **Matriz de frequência**: categorização dos termos associando-os às frequências de ocorrências nos textos.

Ainda sobre o aprendizado de ontologias aliado às técnicas de mineração de texto, Nascimento, Corrêa e Pinho (2019, p. 148-150) reuniram os processos operacionais apoiados em quatro camadas da ontologia e, com base em Schiessl e Brascher (2011), dispuseram as seguintes características:

> **Camada de termos**: Como exemplo de suporte para atender esta necessidade temos as ferramentas de Processamento de

> Linguagem Natural (PLN) atuais, ao qual apresentam resultados satisfatórios na identificação de termos a partir de textos. Essas ferramentas são a combinação de processamentos estatísticos e análises linguísticas. As técnicas mais utilizadas são: parsers para extração de termos, POS tagger (etiquetadores), algoritmos de identificação de palavras-chave, distribuição estatística de termos e testes estatísticos para identificação de termos compostos [...]. Essa estratégia tem como propósito reconhecer as necessidades informacionais mais latentes dos usuários do sistema, neste processo pode-se considerar os métodos de indexação automática.
> **Camada de sinônimos**: As principais técnicas utilizadas para descoberta de sinônimos são de agrupamento (clustering), Latent Semantic Indexing (LSI) a utilização de dicionários especializados de apoio, como tesauros, por exemplo. [...]
> **Camadas de conceito e hierarquização de conceitos**: considera-se como elementos chaves a intensão e extensão do conceito. Para esta camada podem ser inferidas como propriedades de intensão do conceito [...] Para a camada de hierarquização de conceitos, adota-se a taxonomia que pode ser construída por meio da mineração de texto.
> **Camadas de relações e regras**: As principais técnicas se apoiam em análise linguísticas, algoritmos de descoberta de agrupamento e de recorrências de termos. Nesse sentido as regras ou os axiomas são os elementos que dão sentido à formulação de questões de busca, realizando o exercício para validação das relações entre os conceitos. [...] Ainda na camada de relações e regras, citamos a subcamada hiperonímia e hiponímia, em que nesse processo as principais técnicas se apoiam em análise linguísticas, algoritmos de descoberta de agrupamentos hierárquicos e de ocorrência de termos.

A potencialidade do uso de ontologias para atividades estratégicas arremata uma série de pesquisas que se apoiam em sistemas ontológicos para execução automática de atividades operacionais. A esse respeito, podem-se citar a pesquisa intitulada “Ontologia como ferramenta para interligação de dados governamentais abertos: uma análise comparativa” de Rocha, Campos e Costa, que objetivou analisar “[...] a partir de procedimentos metodológicos, a aplicação do uso de ontologias para a interligação de dados governamentais abertos levando em consideração o contexto informacional brasileiro” (ROCHA; CAMPOS; COSTA, 2016, p. 1); e a pesquisa “*Ontologias no processo de indexação automática de documentos digitais*” de Pansani Junior e Ferneda (2016), que teve como objetivo

> [...] propor e desenvolver um método de utilização de ontologias no processo de indexação automática de documentos textuais, fazendo uso da estrutura lógica e conceitual das ontologias de domínio e implementado um método que permite aos sistemas de indexação automática a realização de inferências automáticas, favorecendo uma representação dos documentos mais semântica e abrangente (PANSANI JUNIOR; FERNEDA, 2016, p. 1).

Conforme visto na seção anterior, sobre taxonomias, bem como corroborando com a ideia de que se vive, hoje, numa era do *big data* e suas derivações, a exemplo da mineração de texto, a ontologia pode ser vista, aqui nesta tese, como um sistema de organização do conhecimento, que, apesar de surgir na Ciência da Informação e na Ciência da Computação, em meio tecnológico e digital, perpassou por ricos procedimentos metodológicos, desenvolvidos de acordo com as perspectivas conceituais, processuais e operacionais dos idealizadores.

Neste sentido, de delimitação das perspectivas conceituais e metodológicas quanto ao sistema ontologia, em que se percebeu a forte influência de contributos tecnológicos e digitais, a leitura desta tese segue para a seção de Considerações Finais.

5 CONSIDERAÇÕES FINAIS

Aceitar que a sociedade avança, rapidamente, rumo a uma realidade onde a tecnologia domina parte considerável da rotina de profissionais ligados à informação, é urgente. Tal como o reconhecimento desse avanço no exercício profissional, como o de Bibliotecários e Gestores de Informação por exemplo, insurge-se, ainda, entender que os produtos desenvolvidos, aplicados e operados por esses profissionais acompanham essa velocidade. Nesse sentido, podem-se citar os sistemas de organização do conhecimento: tesauro, taxonomia e ontologia.

Tesauros, taxonomias e ontologias são, entre os demais aqui apresentados, tipos de Sistemas de Organização do Conhecimento que, desde suas primeiras aparições na Ciência da Informação, vêm recebendo atenção de pesquisadores, profissionais e ensaístas da área de Organização e Representação da Informação e do Conhecimento. Essa atenção deve-se, também, ao aperfeiçoamento desses sistemas, em razão do aprimoramento tecnológico e digital.

Sistemas de Organização do Conhecimento, em essência, derivam de aspectos que originam e legitimam múltiplos enredos. Estes são dotados de discussões epistêmicas, históricas, pragmáticas e empíricas.

Seu desenvolvimento, na Ciência da Informação, materializa a ideia da necessidade dialógica com outras áreas do conhecimento, tais como: Linguística; Terminologia; Psicologia; Matemática; Estatística; Tecnologia etc.

Assim como discutido nas linhas introdutórias desta tese, a tecnologia é um elemento-chave que contribuiu e contribui para a linha evolutiva da Ciência da Informação. Em um breve retorno à seção introdutória desta tese, há de se corroborar com Souza (2015) ao compreender que esta ciência é vista como aquela que tem, em sua história, a ampliação e diversificação dos processos de tratamento, organização e recuperação, o que, de tal maneira, acrescentou neste percurso evolutivo, propiciando novos e diversos suportes, processos, tecnologias etc.

A diversificação e evolução desses elementos somam à Ciência da Informação novos ensaios, teorias e práticas, que enriquecem discussões,

desenvolvimento, aplicação e uso dos produtos de organização e representação da informação e do conhecimento, tais como os sistemas aqui estudados.

É imprescindível reiterar que a linha evolutiva da Ciência da Informação não descarta, nem desconsidera, ensaios, teorias e práticas já estabelecidos, mas os enriquece, considerando a vasta diversidade operacional que, por sua vez, conta com uma pluralidade de objetos, instrumentos e meios tecnológicos. Essa pluralidade, no entanto, resulta no aperfeiçoamento de sistemas de organização do conhecimento. Esse aperfeiçoamento foi aqui discutido sob quatro perspectivas apoiadas nos objetivos específicos dessa pesquisa.

Os objetivos específicos, detalhados abaixo, bem como o objetivo geral desta tese desdobraram-se na Metodologia da Problematização e no Arco de Maguerez. Essa intersecção favoreceu o enriquecimento metodológico e serviu como uma trilha científica para concretização dessa pesquisa. Isso porque o método do arco “[...] manteve seu principal objetivo, a saber, propor um caminho para facilitar a busca e construção de conhecimentos a serem implementadas na realidade dos indivíduos [...]” (MAGUEREZ, 2012, p. 13).

O primeiro objetivo específico, da reunião e caraterização dos SOC, refletiu sobre a percepção temporal do conceito dos sistemas e os caracterizou quanto às estruturas e funções. Neste objetivo, explorou-se, para além dos sistemas alvos de investigação desta pesquisa, os sistemas de Classificação; Mapas Conceituais; e Folksonomias. Essa ampliação teórica deve-se à necessidade filosófica de entender sobre a razão da existência desses sistemas e como contribuem para os estudos científicos; ensaios; pesquisas etc.

O segundo objetivo específico implicou no recorte desses sistemas, verticalizando os estudos para tesauro, taxonomia e ontologia. Neste, foram desenhados alguns quadros sobre as perspectivas conceituais acerca desses sistemas. Portanto, de acordo com os extratos consultados, pode-se inferir que o estabelecimento teórico-conceitual desses sistemas se vale do emprego do discurso operacional que, diante do aprimoramento das tecnologias digitais, se estende em práticas que resultam em sistemas mais aperfeiçoados. Esse resultado foi concebido pelas discussões conceituais estabelecidas pelos

autores referenciados no quadro. Percebe-se, também, uma tendência desses autores na modelagem de conceitos estabelecidos e que, para isso, se aportam em um discurso técnico que ora recupera traços oriundos da tecnologia digital ora apresentam novas perspectivas conceituais desses sistemas.

O terceiro objetivo acompanhou o anterior, mas teve foco nos processos metodológicos apresentados em pesquisas sobre um determinado sistema. Aqui, portanto, percebeu-se que os primeiros processos metodológicos, reconhecidos no intervalo das buscas dos extratos bibliográficos consultados, somam às etapas de desenvolvimento de um sistema de elementos que facilitam a construção de SOCs, como, por exemplo: automatização de tarefas para desenvolvimento de ontologias; exploração de arquiteturas digitais orientadas ao usuário na taxonomia; virtualização e navegação flexível nas estruturas conceituais de tesauros.

O quarto, da sistematização dos processos metodológicos para construção de sistemas, buscou elucidar as práticas de planejamento, execução, avaliação e manutenção de tesauro, taxonomia e ontologia. A apresentação da modelagem processual viabiliza, na prática, que esses sistemas possam ser melhores e mais explorados pela comunidade de Ciência da Informação, principalmente por pesquisadores, profissionais e ensaístas da área de Organização e Representação da Informação e do Conhecimento.

Tão logo, ao sistematizar os objetivos acima e como resposta à questão de pesquisa de *como os conceitos e processos dos Sistemas de Organização do Conhecimento (tesauro, taxonomia e ontologia) foram aperfeiçoados em relação ao aprimoramento das tecnologias digitais?,* as ideias resumidas abaixo buscam refletir sobre a integração dos objetivos e da questão problema.

Tesauro, de acordo com os extratos consultados e explorados, não foi o sistema idealizado para ser operado de forma pós-coordenada, em instâncias flexíveis, mas para ser operado de forma mais concisa, estática e menos dinâmica. O aperfeiçoamento desse sistema, por meio do aprimoramento tecnológico e digital, permitiu que as estruturas e relacionamentos conceituais fossem enriquecidas em conceitos e processos. Tais como: as aplicações deste sistema em web semântica, facilitando o processo de descrição e

recuperação da informação; as linguagens adotadas de RDF e XML são, por exemplo, uma das iniciativas percebidas nesta tese. Outros contributos podem ser evidenciados nos tesauros, como a leitura desse sistema em relacionamentos, por exemplo, onde, além de serem compreendidos e determinados por humanos, sejam inteligíveis por máquinas.

Logo, no sentido de contextualizar tesauro nas duas eras acrescentadas na linha evolutiva de SOC (Era da Interoperabilidade e Era Big Data), ressalta-se, conforme os autores explorados, que pesquisas têm se aportado de técnicas e ferramentas de mineração de textos, que se utilizam de tesauros para extração automática de informações. Assim, podem-se exemplificar, em um contexto técnico-operacional, as atividades de indexação automática, a tradução automática, a sumarização de textos etc., onde uma ou mais dessas atividades contribui diretamente na representação e recuperação da informação, facilitando o acesso e uso de informação, contextualização ambiental e tomada de decisão.

Do contrário às primeiras visões de tesauro, a taxonomia e a ontologia nasceram em berço tecnológico. Taxonomia, por exemplo, é um sistema orientado às navegações em sítios da web, por meio de uma estrutura classificatória para organização de domínios do conhecimento. Logo, ela é vista como um sistema de representação de informação e conhecimento que busca, metodologicamente, facilitar a organização, recuperação, representação e acesso à informação em portais, como: portais institucionais; bibliotecas digitais.

Quanto às concepções conceituais e processuais, aprimoradas por tecnologias digitais, podem-se citar múltiplos exemplos que contribuíram para que esse sistema, hoje, seja aplicado de forma mais assertiva e mais orientado aos usuários, como a arquitetura da informação e a experiência dos usuários, que são elementos norteadores de muitas das melhorias recebidas por esses sistemas, pois são responsáveis pela identidade digital do ambiente. As estruturas taxonômicas, nos últimos anos, vêm servindo como guias para desenvolvimento de softwares, sobretudo nos ambientes onde são aplicadas. Elas podem contribuir, para além de acesso e navegação, para *findability* e recuperação inteligente etc.

Tão logo, pode-se revisitar a seção sobre taxonomia e corroborar que a intersecção entre taxonomia e tecnologia: promove orientações de estruturas e páginas web; serve como suporte e ferramenta de organização, representação, navegação e acesso aos conteúdos digitais; permitem recuperar e comunicar informações; e gerencia conteúdo e navegação em um ambiente digital. Assim, a aplicação e uso de taxonomias podem resultar em trilhas de aprendizagens; processos de *onboarding* corporativos; treinamentos educativos e profissionais etc.

Ontologia, vista como um artefato puramente sistêmico e tecnológico, pela Ciência da Computação, influenciou as concepções conceituais e processuais adotadas pela Ciência da Informação, não com atenção maior aos artefatos, mas à promoção de um sistema potencial capaz de organizar, representar e recuperar informações de forma mais contextualizada – um sistema rico e dotado de propriedades técnicas, que são capazes de tornar o processamento dos relacionamentos conceituais inteligíveis por humanos e por máquinas.

Como exemplo de domínio concreto das relações entre áreas do conhecimento na Ciência da Informação, ontologia é um exemplo que reúne estudos com outras ciências como Lógica; Inteligência Artificial; Filosofia; Linguística; e muitas outras áreas científicas em que o sistema se propõe. Essa relação tornou-se rica quando sistemas de recuperação de informação utilizaram-se de ontologias como suporte ao processamento de busca e recuperação da informação, apresentando ao usuário informações contextualizadas no domínio de interesse.

Os quadros conceituais e as modelagens dos processos de tesauro, taxonomia e ontologia, explorados aqui nesta tese, evidenciaram o comportamento desses sistemas na área de Organização e Representação da Informação e do Conhecimento. Assim, tornou-se possível sistematizar concepções de como esses sistemas foram aperfeiçoados em relação ao aprimoramento das tecnologias digitais. Nesse sentido, pela reunião dos objetivos específicos em que este estudo se propôs, e pela questão norteadora da pesquisa, espera-se ter enriquecido as reflexões sobre os sistemas de

organização do conhecimento, quanto aos conceitos e processos, e ter cumprido, satisfatoriamente, o objetivo geral aqui proposto.

Cabe destacar que, apesar de ter-se realizado um estudo verticalizado desses três sistemas, não se pretende negar ou refutar os ensaios, estudos e publicações anteriormente desenvolvidos, mas apresentar à comunidade de Ciência da Informação no Brasil as múltiplas possibilidades conceituais e processuais desses sistemas, de forma a facilitar a compreensão, desenvolvimento, aplicação e uso de tesauros, taxonomias e ontologias. Pode-se, ainda, apontar que outros estudos sobre esses sistemas são necessários, no sentido de mapear, por exemplo, os instrumentos de aplicação e operação desses sistemas e a interoperabilidade entre eles e outros sistemas de organização do conhecimento e sistemas de recuperação da informação.

Ainda que explorados de maneira concisa nesta tese, outros sistemas de organização do conhecimento, como Classificação, Folksonomia e Mapas Conceituais, por exemplo, devem ser alvo de investigações técnico-científicas de forma a esclarecer e enriquecer discussões teóricas e aplicações operacionais, seja na e para a Ciência da Informação ou desta ciência para outras áreas de conhecimento.

Considera-se, portanto, que este documento contribui para que se faça caminhar a área de Organização e Representação da Informação e do Conhecimento, ao que percebe García Gutiérrez por considerar que esta área se encontra em um estado de inércia, “[...] justificado por uma carência de posicionamento ético evidenciado pelo escasso desenvolvimento de métodos, de técnicas, e de ferramentas existentes” (NASCIMENTO, 2018, p. 95) para, assim, cumprir os exercícios em que se propõem esta área.

REFERÊNCIAS

AGANETTE, E. C.; ALMEIDA, M. B.; TEIXEIRA, L. M. D. Criação de definições na construção de ontologias: importância e aplicações na prática médica. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 17., 2016, Salvador. **Anais** [...]. Salvador: Universidade Federal da Bahia, 2016. p. 1-21.

AGANETTE, E. C.; TEIXEIRA, L. M. D. Taxonomias coorporativas: uma proposta de procedimento operacional para construção baseada na teoria e na prática. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 18., 2017, Marília. **Anais** [...]. Marília: Universidade Estadual Paulista, 2017. p. 1-20.

AGANETTE, E.; ALVARENGA., L.; SOUZA, R. R. Elementos constitutivos do conceito de taxonomia. **Informação & Sociedade: estudos**, João Pessoa, v. 20, n. 3, p. 77-93, 2010. Disponível em: https://periodicos.ufpb.br/ojs/index.php/ies/article/view/3994 Acesso em: 02 mar. 2020.

ALBUQUERQUE, M. E. B. C.; GAUDÊNCIO, S. M.; SANTOS, R. F. Reflexões Teóricas em Representação da Informação. *In*: ALBUQUERQUE, M. E. B. C.; MARTIN, G. K.; MOTA, D. A. R. (org.). **Organização e Representação da Informação e do Conhecimento**: intersecções teórico-sociais. João Pessoa: UFPB, 2019. p. 13-24.

ALEXANDER, F. Assessing information taxonomies using epistemology and the sociology of Science. **Journal of Documentation**, v. 68, n. 5, p. 725-743, 2012. Disponível em: https://www.emerald.com/insight/content/doi/10.1108/00220411211256058/full/html Acesso em: 10 fev. 2021.

ALEXANDER, F. Devising a framework for assessing the subjectivity and objectivity of information taxonomy projects. **Journal of Documentation**, v. 70, n. 1, p. 4-24, 2014. Disponível em: https://www.emerald.com/insight/content/doi/10.1108/JD-09-2012-0117/full/html Acesso em: 10 fev. 2021.

ALMEIDA, J. F. V. R.; DIAS, G. A. As dimensões metodológicas da organização do conhecimento. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 20., 2019, Florianópolis. **Anais** [...]. Florianópolis: Universidade Federal de Santana Catarina, 2019. p. 1-22. Disponível em: https://conferencias.ufsc.br/index.php/enancib/2019/paper/view/1259/821 Acesso em: 21 ago. 2020.

ALMEIDA, M. B.; MENDONÇA, F. M.; AGANETTE, E. C. Interfaces entre ontologias e conceitos seminais da Ciência da Informação: em busca de avanços na Organização do Conhecimento. *In*: ENCONTRO NACIONAL DE

PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 14., 2013, Florianópolis. **Anais** [...]. Florianópolis: Universidade Federal de Santa Catarina, 2013. p. 1-23.

ALMEIDA, M. B.; PROIETTI, A. B. F. C.; COELHO, K. C.; ANDRADE, A. Q. A ontologia do sangue: uma inciativa para representação e organização do conhecimento sobre Hematologia e Hemoterapia. **Revista Eletrônica de Comunicação, Informação e Inovação em Saúde**, Rio de Janeiro, v. 7, n. 1, p. 1-12, 2013. Disponível em: https://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/703/1348 Acesso em: 30 abr. 2021.

ARAÚJO, C. A. A. **O que é Ciência da Informação?**. Belo Horizonte: KMA, 2018.

ASSIS, J.; MOURA, M. A. Modelos de mapeamento propostos na Norma ISO 25964: desafios para a consolidação de redes metaconceituais. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 17., 2016, Salvador. **Anais** [...]. Salvador: Universidade Federal da Bahia, 2016. p. 1-16.

BAFFI, M. A. T. **Modalidades de pesquisa**: um estudo introdutório. Petrópolis: Vozes, 2002.

BARDIN, L. **Análise de conteúdo**. São Paulo: Edições 70, 2011.

BEGHTOL, C. Ethical decision-making for knowledge representation and organization systems for global use. **Journal of the American Society for Information Science and Technology**, New York, v. 56, n. 9, p. 903-912, 2005.

BELLUZZO, R. C. B. **Construção de mapas**: desenvolvendo competências em informação e comunicação. 2. ed. Bauru: Cá entre nós, 2007.

BERBEL, N. A. N. **A Metodologia da Problematização com o Arco de Maguerez**: uma reflexão teórica-metodológica. Londrina: EDUEL, 2012.

BERNARDINO, M. C.; SIMOES, M. G. M.; GRACIOSO, L. D. S.; COSTA, L. S. F. Bens culturais rurais: relato de experiência da construção de taxonomia no contexto das fazendas históricas paulistas. *In*: GUIMARÃES, J. A. C.;

DODEBEI, V. (org.). **Organização do conhecimento e diversidade cultural.** Marília: ISKO-Brasil; FUNDEPE, 2015. v. 1. p. 381-391.

BERTIR JUNIOR, D. W.; SOERGEL, D.; LIMA, G. A.; MACULAN, B. C. M. S. Semiautomatização de resalções em tesauros: uma proposta para refinamento de relacionamentos semânticos a partir do agrovoc. **Informação & Informação**, v. 22, n. 3, p. 377-404, 2017. Disponível em http://www.uel.br/revistas/uel/index.php/informacao/article/view/28724 Acesso em: 03 fev. 2021.

BH, PASSINHO. **Passsinho de BH**. Belo Horizonte, 18 ago. 2020. Instagram: @passinhdebh. Disponível em: https://www.instagram.com/p/CEBBuk8pypa/ Acesso em: 07 out. 2020.

BINGHAM, J. **On Jon's desk: Roget's Thesaurus of English Words, without which we wouldn't be celebrating Thesaurus Day**. University of Utah: All U Need, 2017. Disponível em https://openbook.lib.utah.edu/on-jons-desk-rogets-thesaurus-of-english-words-without-which-we-wouldnt-be-celebrating-thesaurus-day/ Acesso em 20 fev. 2021.

BLACKBURN, B. Taxonomy design types. **AIIM E-doc Magazine**, Maryland, USA, v. 20, n. 3, p. 14-16, 2006.

BORDONI, E.; DANSA, S. Árvore de Porfírio. **Fundação CECIERJ**, Rio de Janeiro, 28 jun. 2012, Disponível em https://canal.cecierj.edu.br/recurso/9278. Acesso em 07 out. 2020.

BRAGA, F. R. Extração semiautomática de taxonomia para domínios especializados usando técnicas de mineração de textos. **Ciência da Informação**, Brasília, v. 45, n. 3, p. 175-186, 2016. Disponível em: http://revista.ibict.br/ciinf/article/download/4056/3577 Acesso em: 30 abr. 2021.

BRANDT, M. B. **Etiquetagem e Folksonomia**: uma análise sob a óptica dos processos de organização e recuperação da informação na Web. 2009. 144 f. Dissertação (Mestrado em Ciência da Informação) - Universidade de Brasília, Brasília, 2009. Disponível em: http://repositorio.unb.br/handle/10482/7057. Acesso em: 21 ago. 2020.

BRASCHER, M.; CAFÉ, L. Organização da Informação ou Organização do Conhecimento?. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 9., 2008, São Paulo. **Anais** [...]. São Paulo: Universidade de São Paulo, 2008. p. 1-14. Disponível em: http://enancib.ibict.br/index.php/enancib/ixenancib/paper/viewFile/3016/2142. Acesso em: 21 ago. 2020.

BRAZ, M. I. **Panorama das contribuições da Terminologia para a Organização do Conhecimento**: uma análise das teses e dissertações no Brasil a partir da Teoria da Complexidade. 2020. 205 f. Tese (Doutorado em Ciências da Linguagem) - Universidade Católica de Pernambuco, Recife, 2020.

CAFÉ, L. M. A.; SANTOS, R. N. M.; BARROS, C. M. Os estudos de Gruber e Guarino sobre ontologias na Ciência da Informação e nas Ciências da Computação. **DataGramaZero – Revista de Informação**, Rio de Janeiro v. 16, n. 3, 2015. Disponível em: https://brapci.inf.br/index.php/res/v/8169 Acesso em: 20 abr. 2021.

CAMPOS, L. M.; CAMPOS, M. L. A.; CAMPOS, M. L. M. Diretrizes para definição de domínio no reuso de ontologias biomédicas: uma abordagem baseada na análise do compromisso ontológico. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 11., 2010, Rio de Janeiro.

Anais [...]. Rio de Janeiro: Universidade Federal do Estado do Rio de Janeiro, 2010. p. 1-24.

CAMPOS, M. L. A.; MEDEIROS, J. S. Tesauros conceituais e ontologias de fundamentação: aspectos interdisciplinares na representação de domínios do conhecimento. *In*: CERVANTES, B. M. N. (org.). **Horizontes da organização da informação e do conhecimento**. Londrina: EDUEL, 2012. p. 97-118.

CAMPOS, M. L.; CAMPOS, M. L. A.; SOUZA, J. C. C. E. Estudo de requisitos para escolha de softwares de tesauro. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 18., 2017, Marília. **Anais** [...]. Marília: Universidade Estadual Paulista, 2017. p. 1-20. Disponível em http://enancib.marilia.unesp.br/index.php/XVIII_ENANCIB/ENANCIB/paper/view/453 Acesso em: 21 ago. 2020.

CAPURRO, R. Epistemologia e ciência da informação. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 5., 2003, Belo Horizonte. **Anais** [...]. Belo Horizonte: Universidade Federal de Minas Gerais, 2003. p. 1-21. Disponível em: http://www.capurro.de/enancib_p.htm Acesso em: 21 ago. 2020.

CAPURRO, R.; HJØRLAND, B. O conceito de informação. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 12, n. 1, p. 148-207, 2007. Disponível em: http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/54/47 Acesso em: 21 ago. 2020.

CARIBÉ, R. C. V.; OLIVEIRA, E. S. Avaliação do tesauro do Tribunal de Contas do Distrito Federal a partir dos usuários. **Revista Digital de Biblioteconomia e Ciência da Informação**, São Paulo, v. 13, n. 1, p. 103-119, 2015. Disponível em: https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/1583 Acesso em: 09 jan. 2021.

CARLAN, E. **Sistemas de organização do conhecimento**: uma reflexão no contexto da Ciência da Informação. 2010. 100 f. Dissertação (Mestrado em Ciência da Informação) – Universidade de Brasília, Brasília, 2010.

CARLAN, E.; MEDEIROS, M. B. B. Sistemas de Organização do Conhecimento na visão da Ciência da Informação. **RICI: Revista Ibero-americana de Ciência da Informação**, Brasília, v. 4, n. 2, p. 53-73, 2011. Disponível em: https://doi.org/10.26512/rici.v4.n2.2011.1675 Acesso em: 03 fev. 2021.

CASARIN, H. C. S.; FERREIRA, T. G.; MILANI, S. O. As Crianças como Comunidades de Usuários e os Sistemas de Organização do Conhecimento: uma análise da literatura. *In*: PINHO, F. A.; GUIMARÃES, J. A. C. (org.). **Memória, tecnologia e cultura na organização do conhecimento**. Recife: Editora da UFPE, 2017. p. 361-366.

CHENG, YI-YUN.; LUDASCHER, B. Exploring Geopolitical Realities through Taxonomies: The Case of Taiwan. **Proceedings from North American**

Symposium on Knowledge Organization, v. 7, p. 77-93, 2019. Disponível em: https://journals.lib.washington.edu/index.php/nasko/article/view/15625 Acesso em: 10 fev. 2021.

CONEGLIAN, C. S.; GONÇALVEZ, P. R. A.; SEGUNDO, J. E. S. O profissional da informação na Era do Big Data. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, Brasília, v. 22, n. 50, p. 128-143, 2017. Disponível em: https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2017v22n50p128 Acesso em: 09 mar. 2021.

CONWAY, S.; SLIGAR, C. Building taxonomies. *In*: CONWAY, S.; SLIGAR, C. (org.). **Unlocking knowledge assets**. Redmont: Microsoft Press, 2002. cap. 6.
CORCHO, O.; FERNÁNDEZ-LÓPEZ, M.; GÓMEZ-PÉREZ, A. Methodologies, tools and languages for building ontologies. Where is Their Meeting Point?. **Data & Knowledge Engineering**, Amsterdam, v. 46, n. 1, p. 41-64, 2003.

CÓRDOVA, R. D.; LAZARO, C. J. R. El uso de los sistemas tradicionales de organización del conocimiento en las bibliotecas peruanas. **Biblios (online)**, Peru, n. 46, p. 26-32, 2012. Disponível em: https://biblios.pitt.edu/ojs/index.php/biblios/article/view/38 Acesso em: 21 ago. 2020.

CORDOVIL, V. R. S.; FRANCELIN, M. M. Organização e representações: uso de mapa mental e mapa conceitual. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 19., 2018, Londrina. **Anais** [...]. Londrina: Universidade Estadual de Londrina, 2018. p. 937-956. Disponível em: http://enancib.marilia.unesp.br/index.php/XIX_ENANCIB/xixenancib/paper/view/1301/1502 Acesso em: 21 ago. 2020.

CORRÊA, F.; LACERDA, M. E.; ZIVIANI, F.; FRANÇA, R. S.; RIBEIRO, J. S. A. N. Tecnologias de apoio a Gestão do Conhecimento; uma abstração por conceito, taxonomia e tipologia. **Revista Ibero-Americana de Ciência da Informação**, Brasília, v. 11, n. 2, p. 498-522, 2017. Disponível em: https://periodicos.unb.br/index.php/RICI/article/view/8340/9631 Acesso em: 21 jul. 2020.

CUNHA, J. H. C.; ARAÚJO JÚNIOR, R. H. Taxonomia de distorções contábeis. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, v. 22, n. 49, p. 127-141, 2017. Disponível em: https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2017v22n49p127. Acesso em: 21jun. 2020.

CURRÁS, E. **Ontologias, taxonomia e tesauros em Teoria de Sistemas e Sistemática**. Brasília: Thesaurus, 2010.
DAHLBERG, I. Knowledge organization: a new science? **Knowledge Organization**, Würzburg, v. 33, n. 1, p. 11-19, 2006.

DAHLBERG, I. Teoria do Conceito. **Ciência da Informação**, Rio de Janeiro, v. 7, n. 2, p. 101-107, 1978. Disponível em: http://revista.ibict.br/ciinf/article/download/115/115/ Acesso em: 21 ago. 2020.

DAVANZO, L.; MOREIRA, W. A Teoria do Conceito e a representação da informação arquivística: breves reflexões. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 18., 2017, Marília. **Anais** [...]. Marília: Universidade Estadual Paulista, 2017. p. 1-16. Disponível em http://enancib.marilia.unesp.br/index.php/XVIII_ENANCIB/ENANCIB/paper/viewFile/525/1168 Acesso em: 21 ago. 2020.

DIALLO, G. An effective method of larger scale ontology matching. **Journal of Biomedical Semantics**, [*S.l.*], v. 45, n. 44, p. 1-19, 2014. Disponível em: https://jbiomedsem.biomedcentral.com/articles/10.1186/2041-1480-5-44 Acesso em: 09 mar. 2021.

DIAS, E. W.; NAVES, M. M. **Análise de Assunto**: teoria e prática. Brasília: Briquet de Lemos, 2013.

DÍEZ, M. L. A.; PÉREZ-LEÓN, B.; MARTÍNEZ-GONZÁLEZ, M. M.; BLANCO, D. J. V. Propuesta de representación del tesauro Eurovoc en SKOS para su integración en sistemas de información jurídica. **Scire: representación y organización del conocimiento**, [*S.l.*], v. 16, n. 2, 2010. Acesso em https://ibersid.eu/ojs/index.php/scire/article/view/4015 Acesso em 09 de mar. 2021.

DING, Y. Ontology: the enabler for the semantic web: a review of ontologies with the semantic web in view. **Journal of Information Science**, London, v. 27, n. 6, p. 377-384, 2001.

DZIEKANIAK, G. V. Desenvolvimento de uma ontologia sobre componentes de ontologias. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 15, n. 1, p. 173-184, 2010. Disponível em: https://brapci.inf.br/index.php/res/v/37585 Acesso em: 02 fev. 2021.

FARIAS, M. G. G.; FARIAS, G. B. Aplicação de mapas conceituais como ferramentas didático-pedagógicas na área de recursos e serviços de informação. **Biblios (online)**, Peru, n. 63, p. 13-27, 2016. Disponível em: https://brapci.inf.br/index.php/res/v/69320 Acesso em: 21 ago. 2020.

FELIPE, A. A. C. **O gênero tesauro: um modelo de avaliação linguística**. 2016. 167 f. Tese (Doutorado em Linguística) – Universidade Federal do Rio Grande do Norte, Natal, 2016.

FELIPE, A. A. C.; SILVA NETO, J. G. A dimensão textual do gênero tesauro: uma análise da estruturação composicional e sequencial do tesauro da justiça eleitoral. *In*: ALBUQUERQUE, M. E. B. C.; MARTINS, G. K.; MOTA, D. A. R. (org.). **Organização e representação da informação e do conhecimento**: intersecções teórico-sociais. João Pessoa: Editora UFPB, 2019. p. 29-52.

FERNEDA, E.; DIAS, G. A. OntoSmart: um modelo de recuperação de informação baseado em ontologia. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 22, n. 2, p. 170-187, 2017. Disponível em:

https://www.scielo.br/j/pci/a/TsBd75NfSR99QphrgqpSBWx/?lang=pt Acesso em: 05 fev. 2021.

FERNEDA, E.; DIAS, G. A. Um método de expansão automática de consulta baseada em ontologia. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 14., 2013, Florianópolis. **Anais** [...]. Florianópolis: Universidade Federal de Santa Catarina, 2013. p. 1-19.

FITTERER, R.; METTLER, T.; ROHNER, P.; WINTER, R. A taxonomy for multi-perspective ex ante evaluation of the value of complementary health information systems. **International Journal Healthcare Technology And Management**, [*S.l.*], v. 12, n. 1, p. 45-61, 2011. Disponível em: https://ieeexplore.ieee.org/document/5428385/metrics#metrics. Acesso em: 10 fev. 2021.

FOGL, J. Relations of the concepts 'information' and 'knowledge'. **International Fórum on Information and Documentation**, The Hague, v. 4, n. 1, p. 21-24, 1979.

FUJITA, M. S. L.; ANDRADE E CRUZ, M. C.; PATRÍCIO, B. O. A construção de tesauros na perspectiva dos manuais de indexação. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 18., 2017, Marília. **Anais** [...]. Marília: Universidade Estadual Paulista, 2017, p. 1-26. Disponível em: http://enancib.marilia.unesp.br/index.php/XVIII_ENANCIB/ENANCIB/paper/view/268. Acesso em: 21 ago. 2020.

GABRIEL JUNIOR, R. F.; LAIPELT, R. C. F. Descrição das relações semânticas para aplicação em KOS: uso do tesauro semântico aplicado (THESA). **P2P & Inovação**, Rio de Janeiro, v. 6, n. 1, p. 117-135, 2020. Disponível em: http://revista.ibict.br/p2p/article/view/3815. Acesso em: 09 jan. 2021.

GABRIEL JUNIOR, R. F.; LAIPELT, R. C. F. Thesa: ferramenta para construção de tesauro semântico aplicado interoperável. **P2P & Inovação**, Rio de Janeiro, v. 3, n. 2, p. 124-145, 2017. Disponível em: http://revista.ibict.br/p2p/article/view/3815. Acesso em: 09 jan. 2021.

GAJZER, M. Text and data mining techniques in aspect of knowledge acquisition for decision support system in construction industry. **Technological and Economic Development of Economy**, [*S.l.*], v. 16, n. 2, p. 219-232, 2010.
GILCHRIST, A. **The thesaurus in retrieval**. London: Aslib, 1971.

GONZALES AGUILAR, A. Creación y visualización de tesauros a partir de un vocabulario de técnicas de la información y de la comunicación. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, Santa Catarina, v. 18, n. 37, p. 303-320, 2013. Disponível em: https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2013v18n37p303. Acesso em: 09 mar. 2021.

GREENBERG, J. Introduction: Knowledge Organization Innovation: design and frameworks. **Bulletin of the American Society for Information and Technology**, [*S.l.*], v. 37, n. 4, p. 12-14, 2011. Disponível em: https://asistdl.onlinelibrary.wiley.com/doi/full/10.1002/bult.2011.1720370404. Acesso em: 03 fev. 2021.

GRUBER, T. R. **What is an ontology?**. [*S.l.: s.n*.], 1993. Disponível em: http://www-ksl.stanford.edu/kst/what-is-an-ontology.html. Acesso em: 06 jun. 2021.

GRUBER, T. Toward principles for the design of ontologies used for knowledge sharing. **International Journal of Human-Computer Studies**, [*S.l.*], v. 43, p. 907-928, 1995.

GUARINO, N. Formal ontology, conceptual analysis and knowledge representation. **International Journal of Human Computer Studies**, [*S.l.*], v. 43, n. 5/6, p. 625-640, 1995. Disponível em: http://www2.iath.virginia.edu/time/readings/ontology-semantics-metaphor/ontology-and-knowledge-representation.pdf Acesso: 06 jun. 2021

HERRE, H. Formal Ontology and the foundations of Knowledge Organization. **Knowledge Organization**, Würzburg, v. 40, n. 5, p. 332-339, 2013.

HJORLAND, B. Semantics and knowledge organization. **Annual Review of Information Science and Technology**, [*S.l.*], v. 41, n. 1, p. 367-405, 2007. Disponível em: https://doi-org.ez16.periodicos.capes.gov.br/10.1002/aris.2007.1440410115. Acesso em: 03 fev. 2021.

HJORLAND, B. Theories are Knowledge Organizing Systems (KOS). **Knowledge Organization**, Würzburg, v. 42, n. 2, p. 113-128, 2015.

HJORLAND, B. What is knowledge organization (KO)?. **Knowledge Organization**, Würzburg, v. 35, n. 2/3, p. 86-101, 2008.

HODGE, G. **Systems of Knowledge Organization for Digital Libraries**: beyond traditional authority files. Washington: The Council on Library and Information Resources, 2000. Disponível em: http://clir.wordpress.clir.org/wp-content/uploads/sites/6/pub91.pdf. Acesso em: 02 fev. 2021.

IBEKWE-SANJUAN, F.; BOWKER, G. C. Implications of big data for knowledge organization. **Knowledge Organization**, Würzburg, v. 44, n. 3, p. 187-198, 2017.

JACINTHO, E. M. S. B.; MOREIRO GONZALEZ, J. A. Oferta de emprego: habilidades necessárias para Arquivistas em empresas no Brasil. **Informação & Informação**, Londrina, v. 24, n. 1, p. 424-441, 2019. Disponível em: https://brapci.inf.br/index.php/res/v/110712. Acesso em: 25 set. 2021.

JASPER R., USCHOLD, M. A. Framework for understanding and classifying ontology applications. *In*: INTERNATIONAL JOINT CONFERENCE ON

ARTIFICIAL INTELLIGENCE (IJCAI-99), 1999, Stockholm, Sweden. **Proceedings** [...]. Stockholm, Sweden: [*s.n*.], 1999.

JIMÉNEZ, A. G. Organización del conocimiento para la documentación en periodismo: situación y prospective. **Scire: representación y organización del conocimiento**, [*S.l*.], v. 22, n. 2, p. 21-28, 2016. Disponível em: https://www.ibersid.eu/ojs/index.php/scire/article/view/4327. Acesso em: 01 jan. 2021.

JORENTE, M. J. V.; PADUA, M. C.; SANTAREM SEGUNDO, J. E. Criação de padrões na Web Semântica: perspectivas e desafios. **Em Questão**, Porto Alegre, v. 23, n. 3, p. 157-178, 2017. Disponível em: https://seer.ufrgs.br/EmQuestao/article/view/70466/41289. Acesso em: 21 ago. 2020.

KAULA, P. N. Rethinking on the concepts in the study of classification. **Herald of Library Science**, Varanasi, v. 28, n. 1-2, p. 30-44, 1984.

KORNYSHOVA, E.; DENECKÈRE, R. Decision-Making ontology for Information System Engineering. *In*: INTERNATIONAL CONFERENCE ON CONCEPTUAL MODELING, 29., 2010, Vancouver. **Anais** [...]. Vancouver, Canada: Hal Open Science, 2010. p. 104-117. Disponível em: https://hal-paris1.archives-ouvertes.fr/hal-00662930/document Acesso em: 02 fev 2021.

KROEFF, M. S.; LEONETTI, F. C. Estudo preliminar do tesauro brasileiro de Ciências do Esporte. **Revista ACB**: Biblioteconomia em Santa Catarina, Florianópolis, v. 17, n. 1, p. 76-104, 2012. Disponível em: https://revista.acbsc.org.br/racb/article/view/825. Acesso em: 21 fev. 2021.

KUMAR BL, V.; NIKAM, K. Development of an Information Support System for Yogic Science using Knowledge Organization Systems. **Knowledge Organization**, Würzburg, v. 41, n. 4, p. 288-295, 2014.

LE COADIC, Yves-François. **A ciência da informação**. 2. ed. Brasília: Briquet de Lemos, 2004.

LIMA, G. A. B. de O. Mapa conceitual como ferramenta para organização do conhecimento em sistema de hipertextos e seus aspectos cognitivos. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 9, n. 2, p. 134-145, 2004. Disponível em: http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/355/164. Acesso em: 21 ago. 2020.

LIMA, G. A. B. O. Organização do conhecimento: pesquisa e desenvolvimento. *In*: GUIMARÃES, J. A. C.; DODEBEI, V. (org.). **Organização do conhecimento e diversidade cultural**. Marília: ISKO-Brasil, FUNDEPE, 2015. p. 670-687.

LIMA, G. A. Organização e representação do conhecimento e da informação na web: teorias e técnicas. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 25, número especial, p. 57-97, 2020. Disponível em:

https://periodicos.ufmg.br/index.php/pci/article/view/22283. Acesso em: 05 fev. 2021.

LIMA, G. S.; KROEFF, M. S.; RIBEIRO JUNIOR, D. I. Tesauro afro-brasileiro: uso estratégico para organização e recuperação da informação. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 15., 2014, Belo Horizonte. **Anais** [...]. Belo Horizonte: Universidade Federal de Minas Gerais, 2014. p. 1144-1151.

LIMA, K. Conceito e Organização do Conhecimento segundo Birger Hjorland. **Medium**, São Paulo, 24 nov. 2017, Disponível em: https://medium.com/conceitual/conceito-e-organiza%C3%A7%C3%A3o-do-conhecimento-segundo-birger-hj%C3%B8rland-cdfa6a10256f. Acesso em: 07 out. 2020.

LOPES, P. T. D.; AGANETTE, E. C.; MACULAN, B. C. M. S. Análise da produção de teses e dissertações sobre taxonomias corporativas e facetadas em ciência da informação e ciência da computação. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 19., 2018, Londrina. **Anais** [...]. Londrina: Universidade Estadual de Londrina, 2018. p. 1169-1177. Disponível em: http://enancib.marilia.unesp.br/index.php/XIX_ENANCIB/xixenancib/paper/view/1549. Acesso em: 21 ago. 2020.

LUZ, C. S. O entorno digital da biblioteca *online*: relação da interface e da usabilidade. **Páginas a&b**, Porto, v. 3, n. 9, 2018. Disponível em: https://ojs.letras.up.pt/index.php/paginasaeb/article/view/4563/4274. Acesso em: 30 abr. 2021.

MACULAN, B. C. M. S. **Estudo e aplicação de metodologia para reengenharia de tesauro**: remodelagem do THESAGRO. 2015. 345 f. Tese (Doutorado em Ciência da Informação) – Universidade Federal de Minas Gerais, Belo Horizonte, 2015.

MACULAN, B. C. M. S.; LIMA, G. A. B. O.; PIEROZZI JUNIOR, I. Refinamento de relações em tesauros: reengenharia do Thesagro. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 16., 2015, João Pessoa. **Anais** [...]. João Pessoa: Universidade Federal da Paraíba, 2015. p. 1-22.

MACULAN, B. C.; LIMA, G. A. B. O. Relacionamentos em tesauros: o valor semântico dos verbos. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 19, n. 4, p. 182-201, 2014. Disponível em: http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/2209/1512. Acesso em: 21 ago. 2020.

MAGUEREZ, C. Prefácio. *In*: BERBEL, N. A. N. **A Metodologia da Problematização com o Arco de Maguerez**: uma reflexão teórica-metodológica. Londrina: EDUEL, 2012, p. 9-13.

MAIA, P. C. C.; VASCONCELOS SOBRINHO, M. ; CUNDURÚ, M. T. Terminologia aplicada à produção científica sobre gestão ambiental: diretrizes à elaboração de um microtesauro. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 22, n. 1, p. 80-90, 2017. Disponível em: http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/2610. Acesso em: 21 ago. 2020.

MARQUES, A. M. M. **Utilização pedagógica de mapas mentais e de mapas conceptuais**. 2008. Dissertação (Mestrado em Expressão Gráfica, Cor e Imagem) - Universidade Aberta, Sintra, 2008.

MARTÍNEZ-GONZÁLEZ, M. M.; MUÑOZ NIETO, M. ThesaurusAPI: una API para la manipulación de tesauros. **Scire: representación y organización del conocimiento**, v. 18, n. 1, 2012. Acesso em: https://www.ibersid.eu/ojs/index.php/scire/article/view/3938. Acesso em: 09 mar. 2021.

MAZZOCCHI, F. Knowledge Organization System (KOS): na introductory critical account. **Knowledge Organization**, Würzburg, v. 45, n. 1, p. 54-78, 2018.

MEDEIROS, J. S.; CAMPOS, M. L. A. Tesauros conceituais e ontologies de fundamentação: modelos conceituais para representação de domínios. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 11., 2010, Rio de Janeiro. **Anais** [...]. Rio de Janeiro: Universidade Federal do Estado do Rio de Janeiro, 2010. p. 1-8.

MELLAL, N. **Réalisation de l'interopérabilité sémantique des systèmes, basée sur les ontologies et les flux 194 d'information**. Chambery: Université de Savoie, 2007.

MELO, M. A. F.; BRASCHER, M. Termo, conceito e relações conceituais: um estudo das propostas de Dahlberg e Hjorland. **Ciência da Informação**, Brasília, v. 41, n. 1, p. 67-80, 2014. Disponível em: http://revista.ibict.br/ciinf/article/view/1419/. Acesso em: 21 ago. 2020.

MENDES NETTO, C.; LIMA, G. A. Visualização de ontologias: estudos e perspectivas. **Informação & Sociedade: estudos**, João Pessoa, v. 27, n. 3, p. 59-72, 2017. Disponível em: https://www.academia.edu/37046270/Visualiza%C3%A7%C3%A3o_de_ontologias_estudos_e_perspectivas. Acesso em: 02 mar. 2020.
MENDONÇA, F. M. M.; ALMEIDA, M. B. Ontoforinfoscience: uma metodologia detalhada para construção de ontologias e sua aplicação no domínio da Medicina. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 16., 2015, João Pessoa. **Anais** [...]. João Pessoa: Universidade Federal da Paraíba, 2015. p. 1-18. Disponível em: http://repositorios.questoesemrede.uff.br/repositorios/bitstream/handle/123456789/2791/6.%20ONTOFORINFOSCIENCE.pdf?sequence=1. Acesso em: 21 ago. 2020.

MENDONÇA, F. M.; ALMEIDA, M. B. Princípios metodológicos para desenvolvimento de ontologias: análise das práticas correntes e proposição de melhorias. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 15., 2014, Belo Horizonte. **Anais** [...]. Belo Horizonte: Universidade Federal de Minas Gerais, 2014. p. 430-447.

MIRANDA, L. S.; COELHO, M. H. S.; DIAS, C. C. Avaliação da atualidade temática do thesauro do Direito do Trabalho. *In*: ENCONTRO REGIONAL DE ESTUDANTES DE BIBLIOTECONOMIA, DOCUMENTAÇÃO, GESTÃO E CIÊNCIA DA INFORMAÇÃO DAS REGIÕES SUDESTE, CENTRO-OESTE E SUL, 5., 2018, Belo Horizonte. **Anais** [...]. Belo Horizonte: Universidade Federal de Minas Gerais, 2018. p. 1-12.

MIRANDA, M. L. C. **Organização e representação do conhecimento:** fundamentos teóricos- metodológicos na busca e recuperação da informação em ambientes virtuais. 2005. 354 f. Tese (Doutorado em Ciência da Informação) - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2005.

MORAIS, M. W.; SOUSA, J, L.; RAMALHO, R. A. S. Representação em SKOS de um microtesauro de conhecimentos estratégicos nas organizações. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 24, n. 2, p. 187-194, 2019. Disponível em: http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/4193. Acesso em: 21 ago. 2020.

MOREIRA, M. A. Mapas conceituais como recurso instrucional e curricular. *In*: MOREIRA, M. A. (org.). **A teoria da aprendizagem significativa e sua implantação em sala de aula**. Brasília: Editora da Universidade de Brasília, 2006.

MOREIRA, W.; SANTOS NETO, M. F. La formación del concepto de ontología en la ciencia de la información: un análisis de las revistas Scire e Ibersid. **Scire: representación y organización del conocimiento**, [*S.l.*], v. 20, n. 2, 2014. Acesso em: https://ibersid.eu/ojs/index.php/scire/article/view/4173. Acesso em: 09 mar. 2021.

MOREIRO GONZÁLEZ, J. A. Las relaciones de Asociación em los vocabulários semânticos: ideas sobre su origen, evolución y prospectiva. *In*: ALBUQUERQUE, M. E. B. C.; MARTIN, G. K.; MOTA, D. A. R. (org.). **Organização e Representação da Informação e do Conhecimento**: intersecções teórico-sociais. João Pessoa: UFPB, 2019. p. 53-78.
MOREIRO GONZÁLEZ, J. A. **Linguagens documentárias e vocabulários semânticos para a web**: elementos conceituais. Salvador: EDUFBA, 2011.

MOREIRO GONZÁLEZ, J. A.; CUADRADO SÁNCHEZ, S.; MORATO LARA, J. Mejora de la interoperabilidad semántica para la reutilizácion de contenidos mediante sistemas de organización del conocimiento. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, Santa Catarina, v. 17, n. 33, p. 46-58, 2012. Disponível em:

https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2012v17n33p46/21710 Acesso em: 03 fev. 2021.

MOTA, D. A. R.; KOBASHI, N. Y. Representação e recuperação de informações: contexto, relevância e pertinência. *In*: ALBUQUERQUE, M. E. B. C.; MARTIN, G. K.; MOTA, D. A. R. (org.). **Organização e Representação da Informação e do Conhecimento**: intersecções teórico-sociais. João Pessoa: UFPB, 2019. p. 183-201.

MOZART, F. **Sistemas de Organização do Conhecimento**. Recife, 22 mai. 2020. Instagram: @mozartfelipe. Disponível em: https://www.instagram.com/p/CAfj4BYnRZ_/. Acesso em: 07 out. 2020.

MUCHERONI, M. L.; SILVA, J. F. M.; SANTOS, E. A. Ontologias Spar e uso de metadados FRBR para elaboração de referências. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 19., 2018, Londrina. **Anais** [...]. Londrina: Universidade Estadual de Londrina, 2018. p. 5177-5194. Disponível em: https://repositorio.usp.br/item/002912553. Acesso em: 02 fev. 2021.

NASCIMENTO, F. M. S. **Uso estratégico de ontologia para organização e gestão da informação jurídica**. 2018. 104f. Dissertação (Mestrado em Ciência da Informação) – Universidade Federal de Pernambuco, Recife, 2018 Disponível em: https://repositorio.ufpe.br/handle/123456789/29646. Acesso em: 28 nov. 2020.

NASCIMENTO, F. M. S.; BRAZ, M. I.; PINHO, F. A. Os mapas conceituais para representação da informação no domínio jurídico. **CAJUR - Caderno de Informações Jurídicas**, Brasília, v. 5, n. 1, p. 5-21, 2018. Disponível em: https://brapci.inf.br/index.php/res/v/38061. Acesso em: 21 ago. 2020.

NASCIMENTO, F. M. S.; CORREA, R. F.; PINHO, F. A. Percurso metodológico para construção de Ontologias Jurídicas. **Informação & Sociedade**, João Pessoa, v. 29, n. 4, p. 135-154, 2019. Disponível em: https://periodicos.ufpb.br/ojs2/index.php/ies/article/view/47430/29298. Acesso em: 21 ago. 2020.

NASCIMENTO, F. M. S.; PINHO, F. A. Ontologia na gestão do conhecimento jurídico. **P2P & Inovação**, Rio de Janeiro, v. 4, n. 2, p. 41-52, 2018. Disponível em: http://revista.ibict.br/p2p/article/view/4202/3606. Acesso em: 21 ago. 2020.

NASCIMENTO, F. M. S.; PINHO, F. A. Sistemas de organização do conhecimento: semelhanças e diferenças. **ConCI: Convergências em Ciência da Informação**, Aracajú, v. 2, n. 3, p. 104-122, 2019. Disponível em: https://seer.ufs.br/index.php/conci/article/view/13673/10499. Acesso em: 03 fev. 2021.

NASCIMENTO, F. M. S.; PINHO, F. A. Uso estratégico de ontologia para organização e gestão da informação jurídica. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 19., 2018, Londrina. **Anais** [...].

Londrina: Universidade Estadual de Londrina, 2018. p. 1149-1168. Disponível em: https://brapci.inf.br/index.php/res/v/102148. Acesso em: 03 fev. 2021.

NORTH, K.; PRESSER, N. H. **Reflexões fundamentais para a prática da gestão do conhecimento**. Recife: Néctar, 2011.

NOVAK, J. D. **Aprender criar e utilizar o conhecimento:** mapas conceptuais como ferramentas de facilitação nas escolas e empresas. Lisboa: Plátano Edições Técnicas, 2000.

NOVO, H. F.; A taxonomia enquanto estrutura classificatória: uma aplicação em domínio de conhecimento interdisciplinar. **Ponto de Acesso**, Salvador, v. 4, n. 2, p. 131-156, 2010. Disponível em: https://periodicos.ufba.br/index.php/revistaici/article/view/4103/3409. Acesso em: 02 mar. 2020.

OLIVEIRA, D. A.; ARAUJO, R. F. Construção de linguagens documentárias em sistemas de recuperação da informação: a importância da garantia do usuário. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, Santa Catarina, v. 17, n. 34, p. 17-30, 2012. Disponível em: https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2012v17n34p17. Acesso em: 09 mar. 2021.

ONTORIA, A. *et al*. **Mapas conceptuais**: uma técnica para aprender. 2. ed. Porto: Asa, 1999.

ORREGO, V. N. Terminología clínica: la base para la representación/ categorización de la información en las historias clinicas. **Informação em Pauta**, Fortaleza, v. 3, n. esp., p. 73-94, 2018. Disponível em: http://www.periodicos.ufc.br/informacaoempauta/article/view/39715. Acesso em: 09 jan. 2021.

PANDO, D. A. **Epistemologia da Organização da Informação**: uma análise de sua cientificidade no contexto brasileiro. 2018. 463 f. Tese (Doutorado em Ciência da Informação) - Universidade Estadual Paulista, Marília, 2018.

PANSANI JUNIOR, E. A.; FERNEDA, E. Ontologias no processo de indexação automática de documentos textuais. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 17., 2016, Salvador. **Anais** [...]. Salvador: Universidade Federal da Bahia, 2016, p. 1-20.
PASTOR SANCHÉZ, J. A. Marcado semántico: tecnologías y aplicación para la representación de sistemas de organización del conocimiento en el contexto Linked Open Data. **Scire: representación y organización del conocimiento**, [*S.l.*], v. 19, n. 2, 2013. Acesso em: https://www.ibersid.eu/ojs/index.php/scire/article/view/4112. Acesso em: 09 mar. 2021.

PEÑA, C. M.; REYES, A. A. M.; MONROY, R. A. M. Consideraciones para construir um tesauro sobre leyes político-electorales usando tecnología digital. *In*: DODEBEI, V.; GUIMARÃES, J. A. C. (org.). **Complexidade e Organização**

do Conhecimento: desafios do nosso século. Rio de Janeiro: ISKO-Brasil, FUNDEPE, 2013. p. 110-131. Disponível em: https://isko.org.br/wp-content/uploads/2021/05/Proceedings-ISKO-Brasil-2013.pdf. Acesso em: 09 mar. 2021.

PEREIRA JUNIOR, C. A.; BOTEGA, L. C. Aplicando ontologia como suporte para projeções situacionais. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 20., 2019, Florianópolis. **Anais** [...]. Florianópolis: Universidade Federal de Santana Catarina, 2018. p. 1-9. Disponível em: https://conferencias.ufsc.br/index.php/enancib/2019/paper/view/1427. Acesso em: 11 ago. 2021.

PIEDADE, M. A. R. **Introdução à teoria da classificação**. Rio de Janeiro: Interciência, 1983.

PIETERSE, V.; KOURIER, D. G. Lists, Taxonomies, Lattices, Thesauri and Ontologies: paving a pathway through a terminological jungle. **Knowledge Organization**, Würzburg, v. 41, n. 3, p. 217-229, 2014.

PINHO, F. A. **Aspectos em representação do conhecimento**: em busca do diálogo entre Antonio García Gutiérrez, Michèle Hudon e Clare Beghtol. 2006. 123 f. Dissertação (Mestrado em Ciência da Informação) - Universidade Estadual Paulista, Marília, 2006. Disponível em: https://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/pinho_fa_me_mar.pdf. Acesso em: 21 ago. 2020.

PINHO, F. A.; MILANI, S. O. Organização do conhecimento crítica: reflexões sobre Representação e Homossexualidade. *In*: AMORIM, I. S.; SALES, R. (org.). **Ensaios em organização do conhecimento**. Florianópolis: UDESC, 2021. p. 13-40.

PLAZA-CARVAJA, M.; DUTRA, M.; MACEDO, D. Uma proposta de taxonomia para categorização das falhas no ambiente das cadeias de suprimentos. **Em questão**, Porto Alegre, v. 23, n. 3, p. 179-206, 2017. Disponível em: https://seer.ufrgs.br/index.php/EmQuestao/article/download/71313/42333. Acesso em: 20 mai. 2020.

PONTES, F. V.; LIMA, G. A. B. O.; A organização do conhecimento em ambientes digitais: aplicação da teoria da classificação facetada. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 17, n. 4, p. 18-40, 2012. Disponível em: https://www.scielo.br/j/pci/a/NKD7KNhN3bnrCYSQzjg3Vtn/?lang=pt. Acesso em: 02 mar. 2020.

PORFÍRIO, F. Aristóteles. **Brasil Escola**, 29 ago. 2019. Disponível em: https://brasilescola.uol.com.br/filosofia/aristoteles.htm. Acesso em: 07 out. 2020.

PRATS GARCIA, E. **La evaluación de mapas conceptuales**: un caso práctico. 2013. 197 f. Proyecto de Fin de Máster (Máster en Tecnología Educativa: e-learning y gestión del conocimiento) - Universitat de les Illes Balears, Espanha, 2013.

RAMALHO, R. A. S. Análise do modelo de dados SKOS: sistemas de organização do conhecimento simples para web. **Informação & Tecnologia**, (ITEC), Marília/João Pessoa, v. 2, n. 1, p. 66-79, 2015. Disponível em: https://periodicos.ufpb.br/ojs/index.php/itec/article/view/25995/14680. Acesso em: 03 fev. 2021.

RAMALHO, R. A. S. **Desenvolvimento e utilização de ontologias em Bibliotecas Digitais:** uma proposta de aplicação. 2010. 146 f. Tese (Doutorado em Ciência da Informação) – Universidade Estadual Paulista, Marília, 2010.

RAMALHO, R. A. S.; ZAFALON, Z. R.; OUCHI, M. T. Ontologia baseada nos FRBR: proposta de aplicação em catálogo online. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 15., 2014, Belo Horizonte. **Anais** [...]. Belo Horizonte: Universidade Federal de Minas Gerais, 2014. p. 4132-4150.

RAUTENBERG, S.; TODESCO, L. J.; STEIL, A. V. Uma ontologia para instrumentos da gestão do conhecimento e agentes da engenharia do conhecimento. **Informação & Sociedade: estudos**, João Pessoa, v. 21, n. 1, p. 111-128, 2011. Disponível em: https://periodicos.ufpb.br/ojs2/index.php/ies/article/view/7172. Acesso em: 02 mar. 2020.

ROCHA, L. CAMPOS, M. L. A. O método analítico-conceitual de Dahlberg: aplicação no contexto dos portfólios de projetos da EMBRAPA. **Conhecimento em Ação**, Rio de Janeiro, v. 5, n. 1, p. 22-48, 2020. Disponível em: https://revistas.ufrj.br/index.php/rca/article/view/32196/19798. Acesso em: 21 ago. 2020.

ROCHA, L. L.; CAMPOS, M. L. A.; COSTA, L. C. Ontologia como ferramenta para a interligação de dados governamentais abertos: uma análise comparativa. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 17., 2016, Salvador. **Anais** [...]. Salvador: Universidade Federal da Bahia, 2016. p. 1-20.

RODRIGUES, M. R.; CERVANTES, B. M. N. Análise de assunto e mapas conceituais: semelhanças nos processos. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 20, n. 4, p. 35-56, 2015. Disponível em: https://www.scielo.br/pdf/pci/v20n4/1413-9936-pci-20-04-00035.pdf. Acesso em: 21 ago. 2020.

RODRIGUES, M. R.; CERVANTES, B. M. N. Organização e representação do conhecimento por meio de mapas conceituais. **Ciência da Informação**,

Brasília, v. 41, n. 1, p. 154-169, 2014. Disponível em: http://revista.ibict.br/ciinf/article/view/1425 Acesso em: 21 ago. 2020.

ROQUETA, M. B. Sistemas de Organización del Conocimiento: uma tipología actualizada. **Informação & Informação**, v. 16, n. 3, p. 122-139, 2011. Disponível em: http://dx.doi.org/10.5433/1981-8920.2011v16n2p122. Acesso em: 03 fev. 2021.

ROSATI, I. *et al*. A thesaurus for phytoplankton trait-based approaches: developement and applicability. **Ecological Informatics**, n. 42, p. 129-138, 2017. Disponível em: https://www.sciencedirect.com/science/article/pii/S1574954117302182. Acesso em: 09 jan. 2021.

SABBAG, D. M. A.; GALDINO, R. Tematres e a construção de tesauros: aspectos de aplicabilidade para o profissional da informação. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo, v. 13, n. esp. CBBD 2017, p. 2566-2579, 2017. Disponível em: https://rbbd.febab.org.br/rbbd/article/view/1040/956. Acesso em: 09 jan. 2021.

SALABA, A. Knowledge Organization Requirements in LIS Graduate Programs. *In*: LYKKE, M.; SVARRE, T.; SKOV, M.; MARTÍNEZ-ÁVILA, D. (org.). Knowledge Organization at the Interface*:* Proceedings of the Sixteenth International ISKO Conference, 2020, Aalborg, Denmark. **Ergon-Verlag**. Advances in Knowledge Organization, v. 17, p. 384-393, 2020. Disponível em: https://doi.org/10.5771/9783956507762-384. Acesso em: 15 de mai. 2021.

SALES, L. F.; SAYÃO, L. F. Uma proposta de taxonomia para dados de pesquisa. **Conhecimento em Ação**, Rio de Janeiro, v. 4, n. 1, p. 31-48, 2019. Disponível em: https://revistas.ufrj.br/index.php/rca/article/view/26337. Acesso em: 30 abr. 2021.

SALES, R. A. Relação entre Organização do Conhecimento e Ciência da Informação na Comunidade Científica Brasileira: uma investigação no âmbito da ISKO-Brasil. *In*: GUIMARÃES, J. A. C.; DODEBEI, V. (org.). **Organização do Conhecimento e Diversidade Cultural**. Marília: ISKO-Brasil, FUNDEPE, 2015. p. 73-84.

SALES, R. Para uma organização do conhecimento contemporânea: contribuições de Foucault, Agamben, Deleuze e Guattari. *In*: AMORIM, I. S.; SALES, R. (org.). **Ensaios em organização do conhecimento**. Florianópolis: UDESC, 2021, p. 13-40.

SANTANA, F. **Arquitetura da Informação e o seu propósito**. [*S.l.*]: Coletivo UX, 2017. Disponível em: https://coletivoux.com/arquitetura-de-informa%C3%A7%C3%A3o-e-o-seu-prop%C3%B3sito-29cd278ebdfe. Acesso em: 25 jun. 2020.

SANTOS, F. E.; LIMA, J. S.; ARAÚJO, I. M. Representação e recuperação da informação em bibliotecas universitárias: o tesauro como elo entre acervo,

indexação e usuários da área de Física. **Ciência da Informação em Revista**, Maceió, v. 4, n. 1, p. 43-58, 2017. Disponível em: https://www.seer.ufal.br/index.php/cir/article/view/2828. Acesso em: 09 jan. 2021.

SANTOS, J. C. F. D.; MOREIRA, W.; FUJITA, M. S. L. Tesauro unesp: integração do registro de autoridade para o tematres. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 19., 2018, Londrina. **Anais [...]** Londrina: Universidade Estadual de Londrina, 2018, p. 1110-1128. Disponível em: http://hdl.handle.net/20.500.11959/brapci/103562 Acesso em: 21 ago. 2020.

SANTOS, J. C. F.; CERVANTES, B. M. N.; FUJITA, M. S. L. Tesauro eletrônico: importação no Tematres e disponibilização na web. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 19., 2018, Londrina. **Anais** [...]. Londrina: Universidade Estadual de Londrina, 2018. p. 1110-1128. Disponível em: http://enancib.marilia.unesp.br/index.php/XIX_ENANCIB/xixenancib/paper/view/1366. Acesso em: 21 ago. 2020.

SANTOS, L. C. M.; BRASCHER, M. Recuperação da Informação em acervos digitais de jornais impressos: proposta para modelagem de ontologia no domínio do futebol. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 16., 2015, João Pessoa. **Anais** [...]. João Pessoa: Universidade Federal da Paraíba, 2015. p. 1-8. Disponível em: http://repositorios.questoesemrede.uff.br/repositorios/handle/123456789/2790?show=full. Acesso em: 21 ago. 2020.

SANTOS, L. C. M.; MARTINS, A. P. A. B.; KANEYOA, P. H. O desenvolvimento de taxonomias de setores mercadológicos para uso na inteligência competitiva. **Revista ACB: Biblioteconomia em Santa Catarina**, Florianópolis, v. 21, n. 3, p. 792-806, 2016. Disponível em: https://www.brapci.inf.br/index.php/res/v/680. Acesso em: 30 mar. 2020.

SANTOS, M. T. **Estudo do processo de apropriação da ontologia pela ciência da informação no Brasil**. 2014. 270 f. Dissertação (Mestrado em Ciência da Informação) – Universidade Federal de Pernambuco, Recife, 2014. Disponível em: http://repositorio.ufpe.br/handle/123456789/12945. Acesso em 14 nov. 2019.

SANTOS, M. T.; CORREA, R. F.; LAPA, R. C. Estudos sobre a apropriação da ontologia pela Ciência da Informação. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 14., 2013, Florianópolis. **Anais** [...]. Florianópolis: Universidade Federal de Santa Catarina, 2013. p. 1-7.
SANTOS, R. F.; CORRÊA, R. F. A folksonomia e a representação colaborativa da informação em ambientes digitais. **Tendências da Pesquisa Brasileira em Ciência da Informação**, [*S.l.*], v. 8, n. 1, p. 69-84, 2015. Disponível em: https://brapci.inf.br/index.php/res/v/119561. Acesso em: 21 ago. 2020.

SANTOS, R. F.; NEVES, D. A. B.; SOUZA, E. D. A organização do Conhecimento como domínio de estudo da Ciência da Informação: uma reflexão a partir dos aspectos epistemológicos. *In*: BARROS, T. H. B.; TOGNOLI, N. B. (org.). **Organização do conhecimento responsável**: promovendo sociedades democráticas e inclusivas. Belém: Editora da Universidade Federal do Pará, 2019. p. 95-104.

SCHIESSL, M.; BRASCHER, M. Do texto às ontologias: uma perspectiva para a ciência da informação. **Ciência da Informação**, Brasília, v. 40, n. 2, p. 301-311, 2011. Disponível em: https://brapci.inf.br/index.php/res/download/56021. Acesso em: 30 abr. 2021.

SILVA, A. R. **A dimensão discursiva da organização do conhecimento na Ciência da Informação brasileira**. 2017. 239 f. Tese (Doutorado em Ciência da Informação) - Universidade de Brasília, Brasília, 2017.

SILVA, D. L.; SOUZA, R. R. Modelagem conceitual baseada em ontologias na organização de documentos multimídia. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 17., 2016, Salvador. **Anais** [...]. Salvador: Universidade Federal da Bahia, 2016. p. 1-21.

SILVA, D. **O que é cluster e para que serve?**. [*S.l.*]: Zendesk, 2021. Disponível em: https://www.zendesk.com.br/blog/cluster-o-que-e/. Acesso em: 30 mai. 2021.

SILVA, D. R.; RIBEIRO, C. J. S.; RODRIGUES, J. G. ONTODIP: organização do conhecimento em saúde com uso de ontologias. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 20., 2019, Florianópolis. **Anais** [...]. Florianópolis: Universidade Federal de Santana Catarina, 2019. p. 1-19. Disponível em: https://conferencias.ufsc.br/index.php/enancib/2019/paper/view/1290. Acesso em: 02 fev 2021.

SILVA, P. P.; FREITAS, C.; SILVA, E. M. Desarollo de un software para la detección automática de tópicos en documentos textuales basada en taxonomía. **Ciencias de la Información**, v. 46, n. 3, p. 55-60, 2014. Disponível em: https://www.semanticscholar.org/paper/Desarrollo-de-un-software-para-la-detecci%C3%B3n-de-en-Silva-Freitas/22f370d65d2e72e9cddb9f3df703010c0af9e0ad. Acesso em: 02 mai. 2020.

SIMÕES, M. G. M. *et al*. Entre os seres e os saberes: a identidade ontológica das taxonomias: ciência, método ou produto?. **Ciência da Informação**, Brasília, v. 45, n. 41-56, p. 41-56, 2016. Disponível em: https://www.brapci.inf.br/index.php/res/v/2071. Acesso em: 30 mar. 2020.
SINHA, R. **A cognitive analysis of tagging**. [*S.l.*]: Blog de Rashmi, 2005. Disponível em: https://rashmisinha.com/2005/09/27/a-cognitive-analysis-of-tagging/#more-112. Acesso em: 20 ago. 2020.

SOUSA, M. R. F. *et al*. INFOARCH: uma ontologia para modelar o domínio da arquitetura da informação para web. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 11., 2010, Rio de Janeiro. **Anais** [...]. Rio de Janeiro: Universidade Federal do Estado do Rio de Janeiro, 2010. p. 1-21.

SOUSA, R. T. B.; ARAÚJO JÚNIOR, R. H. A indexação e criação de taxonomias para documentos de arquivo: proposta para expansão do acesso e integração das fontes de informação. **Brazilian Journal of Information Science: Research Trends**, Marília, v. 11, n. 4, p. 47-56, 2017. Disponível em: http://arquivistica.fci.unb.br/wp-content/uploads/tainacan-items/179240/200422/7508-Texto-do-artigo-24221-2-10-20200629.pdf. Acesso em: 21 jun. 2020.

SOUZA, E. D. **A Ciência da Informação**: fundamentos epistêmico-discursivos do campo científico e do objeto de estudo. Maceió: EDUFAL, 2015.

SOUZA, M.; ALMEIDA, J. F. V. R.; SOUZA, R. R. Ontologias para interfaces de sistemas. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 18., 2017, Marília. **Anais** [...]. Marília: Universidade Estadual Paulista, 2017. p. 1-9. Disponível em: https://brapci.inf.br/index.php/res/v/104326. Acesso em: 21 ago. 2020.

SOUZA, R. R. *et al*. Tesauro do Regime Militar Brasileiro. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 15., 2014, Belo Horizonte. **Anais** [...]. Belo Horizonte: UFMG, 2014. p. 1103-1112.

STEIMER, I. S. G.; LUZ, C. S. Taxonomia para comércio eletrônico: diferentes perspectivas em *Front* e *Back End*. **Ciência da Informação em Revista**, Maceió, v. 2, n. 3, p. 3-14, 2015. Disponível em: https://www.brapci.inf.br/index.php/res/v/36268. Acesso em: 08 mar. 2020.

SUÁREZ SÁNCHEZ, A. Sistemas para la organización del conocimiento definición y evolución histórica. **e-Ciencias de la Información**, [*S.l.*], v. 7, n. 2, p. 1-18, 2017. Disponível em: http://eprints.rclis.org/32323/1/Sistemas.pdf. Acesso em: fev. 2021.

TARTAROTTI, R. D. E.; DAL'EVEDOVE, P. R.; FUJITA, M. S. L. Elaboração de tesauro conceitual de política de indexação para bibliotecas universitárias. *In*: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 16., 2015, João Pessoa. **Anais** [...]. João Pessoa: Universidade Federal da Paraíba, 2015. p. 1-19.

THARANI, K. Just KOS! Enriching digital collections with hipertexts to enhance accessibility of Non-Western knowledge materials in libraries. **Knowledge Organization**, Würzburg, v. 47, n. 3, p. 220-230, 2020.

TUDHOPE, D.; NIELSEN, M. L. Introduction to knowledge organization systems and services. **New review of hypermedia and multimedia**, v. 12, n. 1, p. 3-9, 2006. Disponível em: https://doi.org/10.1080/13614560600856433. Acesso em: 15 fev. 2021.

VALENCIA, M. C. P.; VOGEL, M. J. M.; SILVA, V. P. Tesauro de Acervo Espírita: uma revisão de tesauro constituído. **Biblos: Revista do Instituto de Ciências Humanas e da Informação**, Rio Grande do Sul, v. 30, n. 1, p. 102-142, 2016. Disponível em: https://periodicos.furg.br/biblos/article/view/5200. Acesso em: 09 jan. 2021.

VICKERY, B. **On 'knowledge organisation'**. [*S.l.*]: Wayback machine, 2010. Disponível em: https://web.archive.org/web/20100125050134/http://www.lucis.me.uk/knowlorg.htm#start. Acesso em: 09 mar. 2021.

VIEIRA, E. Filosofia: o pensamento político de Platão. **Blog do Professor Edir**, [*S.l.*], 19 mai. 2014. Disponível em: http://professoredirblog.blogspot.com/2014/05/filosofia-o-pensamento-politico-de.html. Acesso em: 07 out. 2020.

VIEIRA, J. M. L.; SANTOS, M. T.; LAPA, R. C. Estudo da construção e aplicação do tesauro na recuperação da informação de teses e dissertações do Programa de Pós-graduação em Desenvolvimento Urbano. **Biblionline**, João Pessoa, n. esp., p. 71-80, 2010.

VIGNOLI, R. G.; SOUTO, D. V. B.; CERVANTES, B. M. N. Sistemas de organização do conhecimento com foco em ontologias e taxonomias. **Informação & Sociedade: estudos**, João Pessoa, v. 23, n. 2, p. 59-72, 2013. Disponível em: https://periodicos.ufpb.br/ojs/index.php/ies/article/view/15160. Acesso em: 03 fev. 2021.

VILELA, P. J.; GOTTSCHALG, C. D. Ontologias: um tipo único de sistema de organização do conhecimento. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 19., 2018, Londrina. **Anais [...]** Londrina: Universidade Estadual de Londrina, 2018, p. 879-899. Disponível em: https://brapci.inf.br/index.php/res/v/102435 Acesso em: 21 ago. 2020.

VÍQUEZ, C. E. Tesauro Especializado en Ciencias del Movimiento Humano (TECMH). **e-Ciencias de la Información**, Costa Rica, v. 8, n. 2, 2018. Disponível em: https://revistas.ucr.ac.cr/index.php/eciencias/article/view/30009 Acesso em: 09 jan. 2021.

VITAL, L. P. **Representação temática de documento arquivístico**: contribuições do modelo conceitual FRSAD. 2017. 191 f. Tese (Doutorado em Ciência da Informação) - Universidade Federal de Santana Catarina, Florianópolis, 2017.

VITAL, L. P.; CAFÉ, L. Proposta para desenvolvimento de taxonomias em portais corporativos. **Perspectivas em Ciência da Informação**, Minas Gerais, v. 16, n. 4, p. 42-54, out./dez. 2011.

VITAL, L. P.; Taxonomia como elemento estruturante em portais corporativos. **Revista Digital de Biblioteconomia e Ciência da Informação**, Campinas, v.

9, n. 2, p. 25-40, 2012. Disponível em: https://periodicos.sbu.unicamp.br/ojs/index.php/rdbci/article/view/1908. Acesso em: 02 mar. 2020.

VOGEL, M. J. M.; KOBASHI, N. Y. Tesauro funcional para organização de arquivos administrativos. **Páginas a&b**, Porto, n. 12, p. 48-62, 2019. Disponível em: https://ojs.letras.up.pt/index.php/paginasaeb/article/view/6370. Acesso em: 30 abr. 2021.

WIKIPÉDIA. Hortus Delicarum. **Wikipedia, The Free Encyclopedia**, 2020b. Disponível em: https://en.wikipedia.org/wiki/Hortus_deliciarum. Acesso em 07 out. 2020.

WIKIPÉDIA. Ingetraut Dahlberg. **Wikipedia, The Free Encyclopedia**, 2020. Disponível em: https://en.wikipedia.org/wiki/Ingetraut_Dahlberg. Acesso em 07 out. 2020.

WU, Y.; YANG, L. Developing a Taxonomy of Semantic Relations in the Oil Spill Domain for Knowledge Discovery. *In*: SMIRAGLIA, R. (ed.). **Proceedings from North American Symposium on Knowledge Organization**, Los Angeles, v. 5, p. 115-130, 2015. Disponível em: https://journals.lib.washington.edu/index.php/nasko/article/view/15185/12651. Acesso em: 10 fev. 2021.

YAMAOKA, E. J. Ontologia para mapeamento da dependência tecnológica de objetos digitais no contexto da curadoria e preservação digital. **AtoZ Novas Práticas em Informação e Conhecimento**, Curitiba, v. 1, n. 2, p. 65-78, 2012. Disponível em: https://revistas.ufpr.br/atoz/article/view/41313. Acesso em: 30 abr. 2021.

YOULIN, Z.; NUNES, J. M. B.; ZHONGHUA, D. Construction and Evolution of a Chinese Information Science and Information Service (CIS&IS) onto-thesauri. **Knowledge Organization**, Würzburg, v. 41, n. 2, p. 131-144, 2014.

ZAHRA, F. M. *et al*. Ferramentas para aprendizagem de ontologia a partir de textos. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 19, n. 1, p. 3-21, 2014. Disponível em: https://www.scielo.br/j/pci/a/3X3rd8Y3HpjdsyYHNxtNWnL/?format=pdf&lang=pt. Acesso em: 15 fev. 2020.

ZENG, M. L. Knowledge Organization Systems (KOS). **Knowledge Organization**, Würzburg, v. 35, n. 2/3, p. 167-182, 2008.