



UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS SOCIAIS E APLICADAS
DEPARTAMENTO DE CIÊNCIA DA INFORMAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO
DOUTORADO EM CIÊNCIA DA INFORMAÇÃO

NATANNA SANTANA DE MORAIS

CONSTRUÇÃO DE TESAuros E ANÁLISE DE DOMÍNIO: PERSPECTIVAS
CONCEITUAIS, TEÓRICAS E METODOLÓGICAS A PARTIR DAS INFLUÊNCIAS
ÉTNICAS NOS LÉXICOS

JOÃO PESSOA
2024

NATANNA SANTANA DE MORAIS

**CONSTRUÇÃO DE TESAUROS E ANÁLISE DE DOMÍNIO: PERSPECTIVAS
CONCEITUAIS, TEÓRICAS E METODOLÓGICAS A PARTIR DAS INFLUÊNCIAS
ÉTNICAS NOS LÉXICOS**

Tese apresentada ao Programa de Pós-Graduação em Ciência da Informação (PPGCI) da Universidade Federal da Paraíba (UFPB) como requisito parcial para a obtenção do título de Doutora em Ciência da Informação.

Área de concentração: Informação, Conhecimento e Sociedade.

Linha de Pesquisa: Organização, Acesso e Uso da Informação.

Orientadora: Profa. Dra. Virginia Bentes Pinto

JOÃO PESSOA

2024

Catalogação na publicação
Seção de Catalogação e Classificação

M827c Moraes, Natanna Santana de.

Construção de tesauros e análise de domínio :
perspectivas conceituais, teóricas e metodológicas a
partir das influências étnicas nos léxicos / Natanna
Santana de Moraes. - João Pessoa, 2024.

257 f. : il.

Orientação: Virgínia Bentes Pinto.
Tese (Doutorado) - UFPB/CCSA.

1. Construção de tesauros. 2. Tesauros -
Metodologias. 3. Análise de domínio. 4. Influências
étnico-raciais nos léxicos. I. Pinto, Virgínia Bentes.
II. Título.

UFPB/BC

CDU 025.43 (043)



UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO

ATA DE DEFESA DE TESE

Defesa nº 103

Ata da Sessão Pública de Defesa de Tese do(a) Doutorando(a) **NATANNA SANTANA DE MORAIS** como requisito para obtenção do grau de Doutor(a) em Ciência da Informação, Área de Concentração em Informação, Conhecimento e Sociedade e com Linha de pesquisa em Organização, Representação e Tecnologias da Informação.

Aos vinte e oito dias do mês de março de dois mil e vinte e quatro (28/03/2024), às nove horas e trinta minutos, no formato híbrido (Presencial/remoto), na Sala de aula 07 do Bloco da Pós-graduação do CCSA/UEPB e Sala virtual do Google Meet, reuniu-se a banca examinadora designada pelo Colegiado do Programa de Pós-Graduação em Ciência da Informação para avaliar o(a) candidato(a) ao Grau de Doutor(a) em Ciência da Informação na Área de Concentração Informação, Conhecimento e Sociedade, o(a) doutorando(a) **NATANNA SANTANA DE MORAIS**. A banca examinadora foi composta pelos (as) professores(as): Dra. Virgínia Bentes Pinto – PPGCI/UEPB (Presidenta/Orientadora), Dra. Gracy Kelli Martins Gonçalves – PPGCI/UEPB (Examinadora interna), Dra. Maria Cleide Rodrigues Bernardino – PPGCI/UEPB (Examinadora interna), Dr. Heliomar Cavati Sobrinho – UFC (Examinador externo), Dr. Raimundo Nonato Macedo dos Santos – UFPE (Examinador externo), Dra. Maria Elizabeth Baltar Carneiro de Albuquerque – PPGCI/UEPB (Suplente Interna) e Dr. Fábio Assis Pinho – UFPE (Suplente Externo). Dando início aos trabalhos, o(a) Professor(a) Dr.(a). Virgínia Bentes Pinto, Presidente(a) da Banca Examinadora, explicou aos presentes a finalidade da sessão e passou a palavra ao(à) discente para que fizesse oralmente a apresentação do trabalho de tese intitulado: “**CONSTRUÇÃO DE TESAuros E ANÁLISE DE DOMÍNIO: PERSPECTIVAS CONCEITUAIS, TEÓRICAS E METODOLÓGICAS A PARTIR DAS INFLUÊNCIAS ÉTNICAS NOS LÉXICOS**”. Após a apresentação, o(a) doutorando(a) foi arguido(a) na forma regimental pelos examinadores. Respondidas todas as arguições, o(a) Professor(a) Dr.(a). Virgínia Bentes Pinto, Presidente(a) da Banca Examinadora, acatou todas as observações da banca e procedeu para o julgamento do trabalho, concluindo por atribuir-lhe o conceito:

(X)Aprovado ()Insuficiente ()Reprovado.

Observações da Banca:
A banca avaliou a tese como excelente. Porém, com algumas observações.
A Banca recomenda a tese para concorrer ao Prêmio ENANCIB, e da UFPB, e CAPES

Proclamados os resultados e encerrados os trabalhos, eu, Professor(a) Dr.(a) Virgínia Bentes Pinto, Presidente da Banca Examinadora, lavrei a presente ata que segue assinada por mim pelos(as) participantes da banca, juntamente com os pareceres de avaliação da tese e da defesa de tese do(a) doutorando(a), devidamente assinados por seus respectivos avaliadores e em formato digital.

João Pessoa, 28 de março de 2024.

gov.br
Documento assinado digitalmente
VIRGINIA BENTES PINTO
Data: 04/04/2024 14:14:54-0300
Verifique em <https://validar.iti.gov.br>

Dra. Virgínia Bentes Pinto
Presidente/ Orientador (a) – PPGCI/UFPB

gov.br
Documento assinado digitalmente
GRACY KELLI MARTINS GONCALVES
Data: 04/04/2024 15:50:16-0300
Verifique em <https://validar.iti.gov.br>

Dra. Gracy Kelli Martins Gonçalves
Examinador (a) Interno (a) – PPGCI/UFPB

gov.br
Documento assinado digitalmente
MARIA CLEIDE RODRIGUES BERNARDINO
Data: 15/04/2024 15:41:16-0300
Verifique em <https://validar.iti.gov.br>

Dra. Maria Cleide Rodrigues Bernardino
Examinador(a) Interno(a) – PPGCI/UFPB

gov.br
Documento assinado digitalmente
HELIOMAR CAVATI SOBRINHO
Data: 13/05/2024 18:58:37-0300
Verifique em <https://validar.iti.gov.br>

Dr. Heliomar Cavati Sobrinho
Examinador (a) Externo (a) – UFC

gov.br
Documento assinado digitalmente
RAIMUNDO NONATO MACEDO DOS SANTOS
Data: 05/04/2024 10:07:10-0300
Verifique em <https://validar.iti.gov.br>

. Dr. Raimundo Nonato Macedo dos Santos
Examinador (a) Externo (a) – UFPE

Dra. Maria Elizabeth Baltar Carneiro de Albuquerque
Suplente Interno(a) –PPGCI/UFPB

Documento assinado digitalmente
 **FABIO ASSIS PINHO**
Data: 04/04/2024 15:43:03-0300
Verifique em <https://validar.iti.gov.br>

Dr. Fábio Assis Pinho
Suplente Externo (a) – UFPE

Documento assinado digitalmente
 **NATANNA SANTANA DE MORAIS**
Data: 14/05/2024 13:09:16-0300
Verifique em <https://validar.iti.gov.br>

Natanna Santana de Moraes
Doutorando(a)

AGRADECIMENTOS

Agradeço a Deus pela vida e a oportunidade de concluir esta pesquisa.

À minha orientadora, professora Virginia Bentes Pinto, pelo aprendizado e as reflexões de valor inestimável, por compartilhar conhecimento, gentileza, respeito.

Aos meus pais, Antônio e Marilúcia, por todo o amor, apoio e incentivo.

Às professoras, Gracy Kelli Martins Gonçalves, Maria Elizabeth Baltar Carneiro de Albuquerque, Cleide Rodrigues Bernardino, aos professores, Heliomar Cavati Sobrinho, Raimundo Nonato Macedo dos Santos e Fábio Assis Pinho por terem aceitado o convite para fazerem parte da banca, agradeço imensamente pelo apoio, disponibilidade e por compartilharem suas contribuições.

À amiga e companheira de doutorado, Ana Isabel Ferreira, pela amizade, pelo apoio e pela força depositada em nossas conversas.

Ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal da Paraíba (PPGCI-UFPB) e aos professores e colaboradores por tanta dedicação e ensinamentos.

Aos colegas do Grupo de Pesquisa de Representação da Informação da Universidade Federal do Ceará (GPRI-UFC).

Às amigas que ajudaram e estiveram sempre me apoiando e me encorajando a concluir essa etapa.

Gratidão!

*Ayvu rapyta rã i iokuaá ma vy ojeupe,
O yvára py mba'ekuaá gui,
O kuaa-ra-ra vy ma
Mborayú rpyta rã oikuaá ojeupe
Yvy oiko eỹ re,
Pytũ yma mbyte re,
Mba'e jekuaá eỹ re,
O kuaa-ra-ra vy ma
Mborayú rapyta rã i oikuaá ojeupe.*

Os fundamentos do ser foram concebidos
Na origem da futura linguagem humana,
Tecida da sabedoria contida em sua própria divindade
E em virtude de sua sabedoria criadora
Concebeu como primeiro fundamento o Amor.
Antes de existir a terra,
Em meio à Noite Primeira,
E antes de ter-se conhecimento das coisas,
O amor era.

Fragmentos de a Criação do Universo, da Terra e do Homem segundo a tradição oral Guarani –
(Jecupé, 2020, p. 43)

RESUMO

Os tesouros são sistemas de organização do conhecimento (*Knowledge Organization System-KOS*) amplamente utilizados para fins de representação e recuperação da informação. Contudo, acredita-se que esses sistemas estão diante de problemáticas referentes à instabilidade nos processos de comunicação, às dispersões léxicas somadas às outras questões envolvendo a não adequação no tratamento da informação e a não introdução de termos originários dos léxicos que herdaram influências étnicas. Outros impasses decorrem da ênfase dada às vertentes epistemológicas hegemônicas (elites científicas, machistas e racistas) que priorizam metodologias e padrões, inviabilizando a construção de tesouros no cenário brasileiro de diversidade étnico-racial. Diante desse cenário, a Análise de Domínio foi introduzida enquanto paradigma social e integrador de metodologias e aspectos teóricos para descobrir e explorar domínios, objetivando o desenvolvimento de estudos na Ciência da Informação e sistemas de organização do conhecimento. Em destaque, a “produção de classificações especiais e tesouros” se configura como uma das onze abordagens da Análise de Domínio. Nessa perspectiva, a presente pesquisa busca responder ao seguinte problema: De que maneira as pesquisas nacionais e internacionais em Análise de Domínio e construção de Tesouros contribuem e ou divergem conceitual, teórico e metodologicamente das questões étnicas-raciais e culturais levantadas a partir das influências étnicas no léxico do português falado no Brasil? Nesse sentido a pesquisa tem como **objetivo geral**: analisar as possíveis contribuições, divergências e/ou aproximações teóricas, conceituais e metodológicas entre a construção de tesouros, a Análise de Domínio e os estudos dos léxicos a partir das influências étnicas brasileiras, considerando as pesquisas em âmbito nacional e internacional. Assim, foram apontados conceitos basilares: etnia, raça, racismo e nação. **Metodologia**: Trata-se de uma Pesquisa Correlacional e documental cujo levantamento se deu na Base de Dados em Ciência da Informação (BRAPCI); na Coleção Scielo Citation Index da Web of Science e na Library, Information Science & Technology Abstracts (LISTA). O referencial recorreu ao histórico dos estudos léxicos e atestou a extinção de maioria das línguas indígenas e afro-brasileiras. Os procedimentos metodológicos conduziram-se conforme a abordagem quanti-qualitativa e a análise de conteúdo. **Resultados**: o corpus da pesquisa identificou 139 documentos, dentre eles artigos e capítulos de livros. **Conclusão**: A partir das leituras e das inferências, foram visualizadas teorias, paradigmas, abordagens, métodos, análises e propostas metodológicas como pontos de convergência, divergência e aproximação entre a Análise de Domínio, a fundamentação teórico e prática dos tesouros e as perspectivas dos estudos dos léxicos em domínios científicos e não científicos. Dentre os aspectos conceituais, destacaram-se os conceitos e questões de gêneros e diversidade sexual, as garantias literária e cultural, o etnoconhecimento, a hospitalidade cultural e a interseccionalidade, além das implicações referentes à colonialidade do poder e à classificação social. Os aportes teóricos interligados aos tesouros e aos estudos de domínios enfatizaram: as Teorias do Conceito e da Classificação Facetada, as teorias epistemológicas e as perspectivas de conhecimento (Empirismo, Positivismo, Racionalismo, Realismo, Pragmatismo, Hermenêutica e Historicismo). As análises bibliométricas e as propostas metodológicas voltadas para os léxicos apareceram enquanto possibilidades viáveis no desenvolvimento de tesouros em diferentes domínios.

Palavras-chave: Análise de Domínio. Construção de tesouros. Influências étnico-raciais nos léxicos. Metodologias em tesouros.

ABSTRACT

THESAURUS CONSTRUCTION AND DOMAIN ANALYSIS: CONCEPTUAL, THEORETICAL AND METHODOLOGICAL PERSPECTIVES FROM ETHNIC INFLUENCES IN LEXICONS

Thesauri are knowledge organization systems (Knowledge Organization System - KOS) widely used for information representation and retrieval purposes. However, it is believed that these systems are faced with problems relating to instability in communication processes, lexical dispersions in addition to other issues involving inadequacy in the processing of information and the non-introduction of terms originating from lexicons that inherited ethnic influences. Other impasses arise from the emphasis given to hegemonic epistemological aspects (scientific, sexist and racist elites) that prioritize methodologies and standards, making the construction of thesauruses unfeasible in the Brazilian scenario of ethnic-racial diversity. Given this scenario, Domain Analysis was introduced as a social paradigm and integrator of methodologies and theoretical aspects to discover and explore domains, aiming to develop studies in Information Science and knowledge organization systems. In particular, the “production of special classifications and thesauruses” is one of the eleven approaches to Domain Analysis. From this perspective, this research seeks to answer the following problem: How does national and international research in Domain Analysis and construction of Thesauruses contribute to and or diverge conceptually, theoretically and methodologically from ethnic-racial and cultural issues raised from ethnic influences in the lexicon of Portuguese spoken in Brazil? In this sense, the research has as its **general objective**: analyze the possible contributions, divergences and/or theoretical, conceptual and methodological approaches between the construction of thesauruses, Domain Analysis and lexical studies based on Brazilian ethnic influences, considering research at national and international levels. Thus, basic concepts were highlighted: ethnicity, race, racism and nation. **Methodology**: This is a Correlational and documentary research whose survey took place in the Information Science Database (BRAPCI); in the Scielo Citation Index Collection of the Web of Science and in the Library, Information Science & Technology Abstracts (LISTA). The framework drew on the history of lexical studies and attested to the extinction of the majority of indigenous and Afro-Brazilian languages. The methodological procedures were conducted according to the quantitative-qualitative approach and content analysis. **Results**: the research corpus identified 139 documents, including articles and book chapters. **Conclusion**: Based on the readings and inferences, theories, paradigms, approaches, methods, analyzes and methodological proposals were viewed as points of convergence, divergence and approximation between Domain Analysis, the theoretical and practical foundation of thesauruses and the perspectives of studies of lexicons in scientific and non-scientific domains. Among the conceptual aspects, the concepts and issues of gender and sexual diversity, literary and cultural guarantees, ethnoknowledge, cultural hospitality and intersectionality stood out, in addition to the implications regarding the coloniality of power and social classification. The theoretical contributions linked to thesauri and domain studies emphasized: Concept and Faceted Classification Theories, epistemological theories and knowledge perspectives (Empiricism, Positivism, Rationalism, Realism, Pragmatism, Hermeneutics and Historicism). Bibliometric analyzes and methodological proposals focused on lexicons appeared as viable possibilities in the development of thesauruses in different domains.

Keywords: Domain Analysis. Thesaurus construction Ethnic-racial influences on lexicons. Thesaurus methodologies.

RESUMEN

CONSTRUCCIÓN DE TESAURO Y ANÁLISIS DE DOMINIO: PERSPECTIVAS CONCEPTUALES, TEÓRICAS Y METODOLÓGICAS A PARTIR DE LAS INFLUENCIAS ÉTNICAS EN LOS LÉXICOS

Los tesauros son sistemas de organización del conocimiento (Knowledge Organization System - KOS) ampliamente utilizados con fines de representación y recuperación de información. Sin embargo, se cree que estos sistemas enfrentan problemas relacionados con la inestabilidad en los procesos de comunicación, dispersiones léxicas, además de otros problemas que involucran insuficiencia en el procesamiento de la información y la no introducción de términos provenientes de léxicos que heredaron influencias étnicas. Otros impasses surgen del énfasis dado a aspectos epistemológicos hegemónicos (élites científicas, sexistas y racistas) que priorizan metodologías y estándares, haciendo inviable la construcción de tesauros en el escenario brasileño de diversidad étnico-racial. Ante este escenario, se introdujo el Análisis de Dominios como paradigma social e integrador de metodologías y aspectos teóricos para descubrir y explorar dominios, con el objetivo de desarrollar estudios en Ciencias de la Información y sistemas de organización del conocimiento. En particular, la “producción de clasificaciones y tesauros especiales” es uno de los once enfoques del Análisis de Dominio. Desde esta perspectiva, esta investigación busca responder al siguiente problema: ¿Cómo las investigaciones nacionales e internacionales en Análisis de Dominio y construcción de Tesauros contribuyen o divergen conceptual, teórica y metodológicamente de las cuestiones étnico-raciales y culturales suscitadas a partir de las influencias étnicas en el? ¿Léxico del portugués hablado en Brasil? En este sentido, la investigación tiene como **objetivo general**: analizar las posibles contribuciones, divergencias y/o aproximaciones teóricas, conceptuales y metodológicas entre la construcción de tesauros, el Análisis de Dominio y los estudios de léxicos basados en influencias étnicas brasileñas, considerando investigaciones en alcance nacional e internacional. Así, se resaltaron conceptos básicos: etnia, raza, racismo y nación. **Metodología**: Se trata de una investigación correlacional y documental cuyo levantamiento se realizó en la Base de Datos en Ciencias de la Información (BRAPCI); en la Colección Scielo Citation Index de Web of Science y en Library, Information Science & Technology Abstracts (LISTA). El marco se basó en la historia de los estudios léxicos y atestiguó la extinción de la mayoría de las lenguas indígenas y afrobrasileñas. Los procedimientos metodológicos se realizaron según el enfoque cuanti-cualitativo y análisis de contenido. **Resultados**: el corpus de la investigación identificó 139 documentos, entre artículos y capítulos de libros. **Conclusión**: A partir de las lecturas e inferencias, se visualizaron teorías, paradigmas, enfoques, métodos, análisis y propuestas metodológicas como puntos de convergencia, divergencia y aproximación entre el Análisis de Dominio, la fundamentación teórica y práctica de los tesauros y las perspectivas de los estudios de léxicos en dominios científicos y no científicos. Entre los aspectos conceptuales, destacaron los conceptos y cuestiones de género y diversidad sexual, garantías literarias y culturales, etnoconocimiento, hospitalidad cultural e interseccionalidad, además de las implicaciones respecto de la colonialidad del poder y la clasificación social. Los aportes teóricos vinculados a los tesauros y estudios de dominio enfatizaron: Teorías de Clasificación Conceptual y Facetada, teorías epistemológicas y perspectivas del conocimiento (Empirismo, Positivismo, Racionalismo, Realismo, Pragmatismo, Hermenéutica e Historicismo). Los análisis bibliométricos y las propuestas metodológicas centradas en léxicos aparecieron como posibilidades viables en el desarrollo de tesauros en diferentes dominios.

Palabras clave: Análisis de Dominio. Construcción de tesauros. Influencias étnico-raciales en los léxicos. Metodologías del tesoro.

LISTA DE ILUTRAÇÕES

Figura 1 –	Categorias do estudo empírico.....	29
Figura 2 –	Ângulo do conceito de acordo com Sartori.....	49
Figura 3 –	Triângulo do conceito, a representação formal de uma unidade de conhecimento.....	50
Figura 4 –	Modelo de construção do conceito.....	51
Figura 5 –	Exemplos de possíveis características.....	53
Figura 6 –	Espiral de desenvolvimento dos assuntos.....	59
Figura 7 –	Mapa mental de facetas para o assunto ‘alimentos’.....	61
Figura 8 –	Exemplos do <i>Thesaurus of Engineering and Scientific Terms (TEST)</i>	65
Figura 9 –	Linha do tempo dos padrões de tesouros de língua inglesa.....	67
Figura 10 –	Exemplos facetas em estrutura construída segundo a abordagem facetada.....	183
Figura 11 –	Exemplo da faceta “kyky” segundo a abordagem facetada.....	185
Figura 12 –	Abordagens da Ciência da Informação em Nascimento (2006)	188
Figura 13 –	Recorte do exemplo da matriz quadrada do total de trabalhos citantes e autores citados da língua e do povo “Tupinambá”.....	205
Figura 14 –	Rede de cocitação de autores com no mínimo 5 citações do descritor “Tupinambá”	206
Figura 15 –	Rede de acoplamento bibliográfico de autores com no mínimo 1 citação e 1 documento do termo descritor Tupinambá.....	208
Figura 16 –	Rede de co-palavras dos 42 artigos recuperados com termo descritor Tupinambá.....	209
Figura 17 –	Síntese dos aspectos conceituais, teóricos e metodológicos a partir da aplicabilidade da Análise de Domínio e da leitura de tópicos étnico-raciais e culturais.....	224
Quadro 1 –	Tipos de Relações entre conceitos.....	54
Quadro 2 –	Conceitos introdutórios apresentados por Ranganathan.....	56
Quadro 3 –	Categorias fundamentais da análise facetada.....	62
Quadro 4 –	Conceitos centrais da perspectiva analítica-sintética facetada.....	63
Quadro 5 –	Comparativo dos padrões por Somers (1981)	69
Quadro 6 –	Registro de vocábulos de origem africana no português brasileiro: século XX.....	123
Quadro 7 –	Aspectos conceituais.....	126
Quadro 8 –	Teorias, paradigmas e epistemologias.....	151
Quadro 9 –	As Teorias do Conceitos e as Teorias do Conhecimento em Hjørland....	170
Quadro 10 –	Abordagens e análises.....	175
Quadro 11 –	Métodos e propostas metodológicas.....	210
Quadro 12 –	Etapas e elementos do Método Etnográfico em Quesada e Quintana (2015)	213
Quadro 13 –	Padrões em construção de tesouros.....	221
Quadro 14 –	Domínios científicos e domínios não científicos.....	224
Tabela 1 –	Estratégias booleanas e resultados.....	30
Tabela 2 –	Autores mais citados conforme a análise de cocitação de autores do termo descritor Tupinambá.....	206

SUMÁRIO

1	INTRODUÇÃO.....	13
1.1	Problematização.....	18
1.2	Justificativa.....	21
1.3	Hipóteses.....	22
1.4	Objetivos.....	22
1.5	Estrutura da tese.....	23
2	PROCEDIMENTOS METODOLÓGICOS.....	25
2.1	Pesquisa empírica.....	30
3	ASPECTOS EVOLUTIVOS E DE CONSTRUÇÃO DE TESAuros.....	34
3.1	Principais Vertentes.....	38
3.2	Teorias fundamentais na construção dos tesauros.....	42
3.2.1	<i>Teoria Geral da Terminologia.....</i>	<i>43</i>
3.2.2	<i>Teoria do Conceito segundo Dahlberg.....</i>	<i>47</i>
3.2.3	<i>Teoria da Classificação Facetada.....</i>	<i>55</i>
3.3	Evolução dos padrões aplicados à construção de tesauros: <i>Thesaurus of Engineering and Scientific Terms</i> a ISO 25964-2.....	64
3.3.1	<i>Construção de Tesauros e as perspectivas de interoperabilidade e da Web Semântica.....</i>	<i>72</i>
3.3.2	<i>Simple Knowledge Organization System (SKOS).....</i>	<i>75</i>
4	ANÁLISE DE DOMÍNIO NA CIÊNCIA DA INFORMAÇÃO.....	78
4.1	Pressupostos da Análise de Domínio.....	79
4.1.1	<i>Dimensão Ontológica</i>	<i>85</i>
4.1.2	<i>Dimensão Epistemológica.....</i>	<i>87</i>
4.1.3	<i>Dimensão Sociológica.....</i>	<i>90</i>
4.2	Domínio, comunidades discursivas e discurso.....	92
5	QUESTÕES E ASPECTOS FUNDAMENTAIS ACERCA DOS LÉXICOS E DAS INFLUÊNCIAS ÉTNICAS.....	99
5.1	Introduzindo os conceitos de etnia, etnicidade, grupos étnicos, raça e nação.....	99
5.2	Aspectos gerais dos estudos dos léxicos.....	109
5.3	O léxico do português no Brasil e suas influências étnicas.....	116
5.3.1	<i>Influências étnicas indígenas do Brasil.....</i>	<i>117</i>
5.3.2	<i>Influências étnicas africanas: resistência e silenciamento da afro-brasilidade.....</i>	<i>120</i>
6	ANÁLISE E DISCUSSÃO.....	126
6.1	Aspectos conceituais.....	126
6.2	Aspectos teóricos.....	150
6.3	Aspectos metodológicos.....	175
6.4	Síntese.....	224
6.5	Verificação das Hipóteses.....	225
7	CONCLUSÃO.....	229

REFERÊNCIAS.....	232
APÊNDICE A - FICHA DE REGISTRO DOCUMENTAL.....	250
APÊNDICE B – DOCUMENTOS DA PESQUISA EMPÍRICA.....	251

1 INTRODUÇÃO

O presente trabalho iniciou com reflexões e questionamentos acerca da ausência de vocabulários e princípios que pudessem orientar a representação da informação e de conhecimentos registrados em documentos referentes às temáticas étnicas e culturais no Brasil. Essas temáticas estão presentes em acervos bibliotecas comunitárias, públicas e universitárias, principalmente, aquelas que integram trabalhos de conclusão de curso e livros que descrevem memórias e trajetórias históricas de grupos e lugares. Nessa perspectiva, os tesouros se destacam como vocabulários estruturados para auxiliar no tratamento desses documentos, considerando que o seu contexto de aplicabilidade envolve as palavras e os vocábulos advindos do léxico do público-alvo ou do assunto que busca representar.

De acordo com a International Organization for Standardization - ISO (2011, p. 12, tradução nossa) o tesouro é conceituado como sendo um “vocabulário controlado e estruturado no qual conceitos são representados por termos, organizados para que as relações entre conceitos sejam explícitas e os termos preferidos estejam acompanhados por entradas para sinônimos ou quase-sinônimos”. Colaborando nesse entendimento, Soergel (2005) considera o tesouro como uma estrutura capaz de gerenciar as complexidades da terminologia na disposição de relações conceituais utilizadas para especificar descritores autorizados à indexação e à pesquisa. Em outra definição, os tesouros são linguagens documentárias, constituindo-se como “[...] sistemas extremamente complexos, e apresentam, por isso mesmo, problemas de ordem econômica para sua construção e utilização” (Lara, 1999, p. 16).

Um tesouro ao contemplar a terminologia de especialidade dos usuários e a realidade da instituição, incorpora aspectos característicos do contexto do qual se insere. A possibilidade de sua atualização pode ser mais rápida do que os sistemas de classificações bibliográficas (Lara, 1999), a exemplo Classificação Decimal Universal (CDU) e a Classificação Decimal de Dewey (CDD).

Seus parâmetros de atualização e de construção se sustentam em teorias, princípios, normas, abordagens, métodos e diretrizes, juntamente com a aplicabilidade de dispositivos, softwares, aplicativos, padrões e formatos que compõem os sistemas, do mesmo modo se efetivam nos sistemas intranet e internet, bancos e bases de dados. Contudo, há desafios na sua elaboração quanto as dificuldades elencadas durante sua estruturação e manutenção: o alto custo, o longo período de elaboração, o nível de especialização e o grau de formação e interação entre profissionais de inúmeras competências (linguísticas, terminológicas, tecnológicas, entre outras).

O foco da sua elaboração está direcionado à cobertura de um dado domínio, especialidades ou temática, ainda que, em alguns casos tenha a pretensão de se debruçar sobre algumas das áreas do conhecimento, como por exemplo, os tesouros do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), da Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO) e do Thesaurus Science and Technology Policies Information Exchange System (SPINES) desenvolvido pela UNESCO. Cavati Sobrinho (2014) destaca nesse limiar, a necessidade de delimitar o domínio como parte das etapas e do desenvolvimento metodológico de construção de tesouros em específico no domínio da Economia.

Precisamente, conforme Zeng (2008), os tesouros auxiliam os sistemas de classificação e organização do conhecimento, ou seja outros vocabulários controlados, ontologias e taxonomias, empregadas nos sistemas de recuperação da informação. E, por sua vez, são utilizados na arquitetura de informação de *sites* e ambientes da *web* (Rosenfeld; Morville; Arango, 2015). Isso se verifica, considerando suas funções direcionadas à recuperação, ao tratamento e à adequação de linguagens naturais diante dos problemas causados pela inacessibilidade à interpretação e ao entendimento conceitual, lexical e terminológico.

Os tesouros continuaram sendo utilizados e readequados às transformações das últimas seis décadas. Na realidade brasileira, por exemplo, eles estão presentes em bibliotecas, museus, arquivos e, principalmente, em instituições e setores organizacionais que tratam e são encarregados da representação e recuperação da informação e, também, nos acervos documentais de natureza física e digital.

Perante essas reflexões, é necessário analisar nos discursos um sistema nocional, sistema esse que contextualiza as origens e os usos propriamente ditos de seus termos, para além da descoberta do domínio e sua importância no planejamento dos tesouros. Retomamos a exemplo Cervantes (2009, p. 17) que ao propor um modelo metodológico situa os tesouros com a fundamentação da Análise de Domínio: “em razão de nossa pesquisa versar sobre a construção de tesouros em sistemas de informação especializados, no âmbito das abordagens destacadas por Hjørland (2002a) interessa-nos particularmente a segunda abordagem: construção de classificações especiais e tesouros, a qual contempla a construção de tesouros”. Nesse sentido, concordamos com o enfoque dado para os domínios e suas orientações sociocognitivas e socioculturais no âmbito da Organização do Conhecimento (Hjørland; Albrechtsen, 1995; Hjørland, 2002a, 2004a, 2017; Smiraglia, 2012, 2014, 2015a; López-Huertas, 2015)

É nessa perspectiva, que percebemos, empiricamente, a possibilidade de um diálogo com a Análise de Domínio (AD) e a construção de tesouros, bem como outros objetos de

estudos da CI. Isto porque a AD vislumbra componentes e dimensões de um determinado domínio criado com base no conhecimento compartilhado por uma comunidade (Hjørland; Albrechtsen, 1995; Hjørland, 2002a; 2017; Smiraglia, 2015). Nesse contexto, percebe-se também o conhecimento conforme as interações sociais e todos os seus componentes de acordo com a investigação e a descoberta de domínios. Essa noção modificou a centralidade dos estudos e pesquisas, deixando de lado abordagens fixadas em assuntos e documentos, e paradigmas, como o físico e aqueles com o viés positivista.

Diante dessas ponderações, os paradigmas – Social, Cognitivo e Físico- da Ciência da Informação (Capurro; Hjørland, 2007) detêm um papel fundamental nos estudos e pesquisas que buscam investigar e propor soluções e análises para problemáticas relacionadas ao tratamento, representação, organização e acesso à informação, pois se encontram com epistemologias e vertentes científicas capazes de modificar os caminhos a serem percorridos pelos pesquisadores e profissionais.

A partir da perspectiva social da AD, torna-se viável compreender que a construção da ciência se institucionaliza, entre outras coisas, por meio da estruturação epistemológica e terminológica relacionada ao domínio e ao objeto de estudo. Entretanto, além dessa (intra) terminologia tem que se pensar que ela será utilizada em um processo de comunicação que se efetiva para além do âmbito das instituições científicas e tecnológicas.

Tal fato, demanda outros olhares ou reflexões diante dos contextos, indivíduos e elementos que interagem diretamente na compreensão dos domínios. Assim, é possível identificar que as temáticas e os fenômenos do âmbito social progridem em conjunto com as comunidades e os grupos sociais dos quais integram e interagem direta e indiretamente com essas instituições e seus processos, uma vez que a colaboração dos participantes é imprescindível para transcorrer o ciclo científico.

A efetividade dos contributos esboçados na Ciência da Informação (CI) para a construção de tesouros esteve estabelecida paralelamente a consolidação das terminologias, essas essenciais para a confirmação dos discursos e a redução de interferências nos processos de comunicação técnico-científicas. Visam, ainda, acompanhar as tendências dos contextos técnicos, científicos e outros instrumentos de organização e representação da informação e do conhecimento.

Essa ciência acompanhou e se desenvolveu, promovendo, dentre outras contribuições, a elaboração de sistemas responsáveis pela organização e representação da informação e do conhecimento, antes pautados nas linguagens de classificação e nos cabeçalhos de assuntos. A inovação nessa temática, também, desperta com o surgimento dos tesouros.

Esses se destacam na lista dos vocabulários controlados estruturados alinhados aos projetos e tendências aplicadas na composição das linguagens de especialidades.

A CI se estabeleceu a priori enquanto ciência interdisciplinar e possui uma composição teórica e prática para observar contextos humanos e não humanos perante o uso, a produção, o tratamento e a disseminação da informação. Também se tornou uma ciência imprescindível nas discussões e estudos sobre o conhecimento registrado e sua relação intrínseca com a informação. A perspectiva conceitual oriunda de Borko (1968, p. 2, tradução nossa) esclarece a sua dimensão científica na qual assevera que a:

Ciência da Informação é a disciplina que investiga as propriedades e o comportamento informacional, as forças que governam os fluxos de informação, e os significados do processamento da informação, visando à acessibilidade e a usabilidade ótima. A Ciência da Informação está preocupada com o corpo de conhecimentos relacionados à origem, coleção, organização, armazenamento, recuperação, interpretação, transmissão, transformação, e utilização da informação. Isto inclui a pesquisa sobre a representação da informação em ambos os sistemas, tanto naturais quanto artificiais, o uso de códigos para a transmissão eficiente da mensagem, bem como o estudo do processamento e de técnicas aplicadas aos computadores e seus sistemas de programação.

Conforme pode ser observado nesse conceito, desde o seu nascedouro, a CI não somente se prende ao objeto informacional, mas a consolidação da área perante a formação de uma cultura que incube a informação e os suportes do conhecimento como algo inseparável e intrinsecamente ligados. Para tanto, é necessário compreender aspectos conceituais acerca da informação e conhecimento, antes do aprofundamento das atividades de planejamento e aplicabilidade dos tesouros.

Para tanto, Fogl (1979) indica que a informação é uma unidade composta pelos aspectos do conhecimento registrado (conteúdo da informação); pela linguagem percebida como expressão; e o suporte (objetos materiais ou energia). Essa unidade pode ser investigada de acordo com seus aspectos semânticos, pragmáticos e em relação ao método de fixação do conhecimento, ao juízo de valor e com base no suporte material (Bräscher; Café, 2008).

Nesse cenário, Dahlberg (2006) destaca a fundação da *International Society for Knowledge Organization* (ISKO,) ou Sociedade Internacional para a Organização do Conhecimento, em 1989, cujos membros fundadores acataram o termo “Organização do Conhecimento” (OC) utilizado pelo bibliotecário americano Henry Evelyn Bliss, e, que mais tarde em 1993, o periódico *International Classification* introduziu esse conceito para compreender os objetos e as atividades da teoria conceitual da classificação, indexação e representação do conhecimento.

Apesar de questionar a Organização do Conhecimento (OC ou do inglês KO, *Knowledge Organization*) enquanto uma nova ciência, Dahlberg (1993) considera essa, de forma mais clara, como sendo uma disciplina inserida do campo científico da CI e que inclui fundamentos teóricos necessários para as investigações acerca dos sistemas conceituais e do conhecimento, especialmente identificando alguns trabalhos teóricos como as seguintes contribuições de:

- a) Ranganathan - o método de categorização (análise de facetas) e combinações de conceitos;
- b) Otlet e La Fontaine - a revisão e ampliação da Classificação Decimal de Dewey;
- c) Wüster – a constituição de sistemas com base nas duas formas: as hierárquicas e as de sistemas conceituais usando a lógica.

Contudo, a noção é mais elucidada por Esteban-Navarro e García-Marco (1995) entende que a Organização do Conhecimento é uma disciplina dedicada ao estudo e ao desenvolvimento de uma série de fundamentos e técnicas de planejamento, construção, gestão, uso e avaliação de sistemas de descrição, além das atividades catalogação, organização, classificação, armazenamento, comunicação e recuperação, objetivando preservar e transmitir seu conhecimento e suas ações, com base em seu conteúdo, a fim de garantir sua conversão em informação capaz de gerar novos conhecimentos.

A partir dessa explicação, a Organização do Conhecimento é um segmento disciplinar e científico da pragmática e da epistemologia da CI que objetiva a investigação de fundamentos, metodologias e técnicas para gerir, aplicar e mediar estruturas do conhecimento tanto científicas, quanto sociais. A pragmática de suas perspectivas metodológicas e teóricas pode ser vista nas unidades de documentação (Bibliotecas, arquivos, museus) e bancos de dados, a partir da atuação de especialistas que aplicam seus conhecimentos especiais, ou em atividades como descrição, indexação e classificação de documentos e “instituições de memória” (Hjørland, 2008).

Em específico, OC tem seu foco nos Sistemas de Organização do Conhecimento (SOCs, ou KOS - *Knowledge Organization System*) como forma de canalizar seus esforços para estudar de que forma o conhecimento registrado é estruturado e utilizado para seus diferentes propósitos. Dentre os tipos de Sistemas de classificação, listas de cabeçalhos de assuntos, tesouros, ontologias e taxonomias (Zeng, 2008; Carlan; Bräscher, 2011; Hjørland, 2016).

Ao absorver, analisar e transcrever a informação e o conhecimento nas esferas sociais, que uma ciência como a da Informação remodela e reimprime seus estudos de acordo com suas linguagens e seus discursos seja de forma positiva apresentando ou relatando aspectos socioculturais e históricos, seja de forma negativa ajudando no silenciamento e na extinção de muitas identidades e elementos socioculturais e étnicos.

No caso da CI essa forma negativa pode ser estabelecida por meio do controle, da redução, do privilégio, da censura, entre outros posicionamentos que permitem a hegemonia de grupos e ideologias presentes em sistemas de informação e de conhecimento em detrimento de outras comunidades e povos originários brasileiros e africanos que são penalizados constantemente com o apagamento de sua cultura e a retirada de seus direitos.

Conforme destacado por Cardona (2020), a CI e a Biblioteconomia da América Latina não estão isentas, já que suas epistemologias estão localizadas e diante do processo subordinação em relação as ciências modernas oriundas da produção de conhecimentos hegemônicos como o anglo-saxão e o europeu. Essa subordinação é identificação com uma relação de superordenação que impõe uma dominação durante o estabelecimento e disseminação de qualquer tipo de conhecimento.

É nesse ponto, que emergem as inquietações aqui relatadas: a Ciência da Informação e a Organização do Conhecimento detêm participações fundamentais nas contribuições científicas destinadas a atender demandas sociais do tipo acessibilidade facilitada com o auxílio da representação e organização de documentos que registram os diferentes elementos culturais e simbólicos das etnias indígenas e afro-brasileiras. Podemos citar outros tópicos relevantes para essas demandas: usos e necessidades informacionais; representações sociais, identidades de grupos e costumes presentes nos registros e acervos físicos ou digitais portadores de informação. E no caso específico dos tesouros, atribuem-se funções importantes para o funcionamento e a eficácia direcionadas aos sistemas e portais governamentais e de assistência cidadã com toda a totalidade de componentes e vocabulário correspondentes e essenciais para mediar a informação para esses públicos.

1.1 Problemática

Ao tratar da construção de tesouros, é imprescindível e relevante reconhecer os alicerces teóricos e metodológicos de sua elaboração, avaliação e atualização. Para tanto, seguem perspectivas epistemológicas que, assim como em outros domínios, privilegiam

comunidades científicas e discursos hegemônicos destoantes e, em alguns casos, antagônicos das dimensões sociocultural, política, étnica e histórica.

Essas questões podem ser conferidas na problemática dos enfoques (*Bias*) negativos e neutros (Olson, 2002, 2011; Colombo, Barité, 2015) e da decolonização na Organização do Conhecimento (Garcez; Sales, 2021); na violência epistêmica sofrida pelas abordagens e subjetividades decoloniais¹ no contexto latino-americano (Zamudio, 2018; Cardona, Silva, 2020); nas epistemologias negras e as relações raciais (Barroso *et al.*, 2019; Silva, 2020; Silva; Saldanha, 2019), nas questões raciais e no protagonismo das mulheres (Silva; Romeiro, 2019). Ou seja, há disparidades enfrentadas na aplicabilidade desse instrumento, pois conforme já mencionado, em algumas experiências, não são considerados aspectos históricos, geográficos e socioculturais das comunidades de usuários.

Tomamos como exemplo, a sobreposição de vertentes que priorizam e exaltam termos racistas, ou até mesmo o predomínio de conceitos e termos das religiões cristãs em detrimento do apagamento daqueles que se referem às religiões de matrizes africanas. Outros obstáculos podem ser verificados conforme a utilização de normas e diretrizes que inviabilizam a introdução de termos originários de comunidades étnicas e sociais, afrodescendentes, indígenas e a comunidade LGBTQIAPN+ (Lésbicas, Gays, Bissexuais, Transgêneros, Travestis, *Queer*, Intersexos, Assexual, Pansexual, Não-binário e outros).

Esses exemplos evidenciam os vocabulários e palavras adotadas em muitos discursos científicos de trabalhos e livros acerca das temáticas protagonizadas por esses grupos, assim o trabalho teórico e prático absorve e dificulta a integralização de novas epistemologias, perspectivas de igualdade, acessibilidade e diversidade de conhecimentos, discursos e identidades. Na mesma linha de raciocínio, Martins e Côrtes (2019, p. 167-168) defendem que "entre os fenômenos sociais que definem a organização e representação do conhecimento, a linguagem encontra-se como fenômeno multiforme, interdisciplinar e objeto indispensável para estudos em ORIC [Organização e representação da informação e do conhecimento]". Logo, as discussões e as atualizações em ORIC devem atender as adequações dos sistemas de

¹ As abordagens e epistemologias decoloniais são originárias das implicações da ideia de decolonialidade. Bernardino-Costa, Maldonado-Torres e Grosfoguel (2018) apresentam o conceito de decolonialidade enquanto projeto acadêmico-político, do qual, dentre suas vantagens, está a capacidade de esclarecer e sistematizar o somatório daquilo que é elucidado a partir das reações frente ao colonialidade do poder e sua dominação colonial, cuja dimensão política é caracterizada pelas lutas políticas de resistência e reexistência das populações afrodiáspóricas e africanas, indígenas e do terceiro mundo. Esse conceito e seus desdobramentos viabilizam o entendimento do reconhecimento das múltiplas experiências e epistemologias subordinadas e destituídas pelo projeto do colonialismo europeu.

organização condizentes com as novas políticas, tanto sociais quanto linguísticas, em prol da diversidade e dos grupos socialmente minorizados.

Ainda a esse respeito, Pinho e Guimarães (2011, p.143) argumentam que “o instrumento e os produtos da representação do conhecimento não são neutros, uma vez que seus idealizadores impõem uma visão de mundo particular, refletindo posições ideológicas e políticas”. Nesse sentido, qualquer indício de atuação comprometida com a ética e a bioética precisam ser reintroduzidos nos estudos e nas ações, já que as representações se colocam em meio a subjetividade do profissional e as influências ideológicas.

Além disso, as dificuldades de tratar a linguagem natural e a linguagem terminológica estão presentes no tratamento dos aspectos sociais, culturais e étnicos mediante as pesquisas nos campos das ciências sociais e humanidades. Uma questão central pode ser destacada nesse tópico: a necessidade de compreender a dimensão sociocultural perante a dicotomia entre conhecimento científico e o conhecimento advindo do âmbito social, seja saberes populares, seja quaisquer manifestações intelectuais e artísticas, como a literatura e o conhecimento religioso.

Citamos como exemplo desse empecilho, uma pesquisa realizada em 2022 no catálogo da Biblioteca da Universidade Federal do Ceará (UFC) que durante a busca pelo assunto “indígena” (com o filtro especificado em assunto) foram detectados 34 títulos, contemplando somente os seguintes tipos documentais e suas respectivas quantidades: Dissertações (19); Monografias (6); e Teses (9). Entretanto, o índice de termos descritores relacionados a essas publicações foi limitado, consistindo em um número reduzido de termos relativos ao assunto indígena e algumas publicações foram indexadas com o termo “índio”. O termo “índio” se remete a uma forma pejorativa e representante de uma noção do ponto de vista histórico da colonização.

Ainda nessa pesquisa, foram detectados 97 termos descritores, dentre os quais: Desenvolvimento sustentável; Direito à saúde; Discriminação; Ecologia agrícola; e Economia social. Parte desses termos não contemplam aspectos étnicos e culturais, ou suas respectivas características, que facilitam o descobrimento do pesquisador diante da rica literatura acerca do acervo de trabalhos dos cursos de Licenciatura Intercultural Indígena das Etnias Pitaguary, Tapeba, Kanindé, Jenipapo-Kanindé e Anacé da UFC. Talvez esse fato tenha acontecido em razão do uso de uma Linguagem controlada de acordo com o Catálogo de Autoridades da *Library of Congress*, sem que tenham sido observadas as evidências que atestam o uso dos termos (assunto) conforme as características locais do acervo, do público e do domínio em questão.

Assim, há necessidade de adequar os termos correspondentes à abordagem sociocultural, às comunidades discursivas (Hjørland, 2002a) e as garantias norteadoras dos procedimentos de representação (Dias, 2015), viabilizando o tratamento das linguagens naturais – que incluem as palavras de determinado léxico – e do processo de tradução que contempla o sistema nocional e a natureza terminologia de cada disciplina e domínio (Dodebei, 2002).

Nesse contexto, a Análise de Domínio está centrada em diálogos que permitem, de acordo com o seu aporte epistêmico e social, contribuir com os estudos dos sistemas de organização do conhecimento, em específico contribuir com a produção de Classificações Especiais e Tesauros (Hjørland, 2002a).

Reforçando, ainda, a aplicabilidade de métodos e princípios metodológicos, dentre eles os Estudos Métricos da Informação (EMI) e as análises bibliométricas na compreensão e visualização de domínios (Glänzel, 2003; Zhou; Chen; Wang, 2009; Oliveira, Grácio, 2013; Grácio, 2016, 2020), como demonstrou Rees-Potter em 1989 e 1991 ao descrever a experiência da análise de citação e co-citação para construção e manutenção de tesauros segundo as premissas da AD, na qual os termos manipulados foram extraídos de forma semiautomática a partir dos contextos de citação dos artigos e de acordo com citações em áreas de especialidade (Rees-Potter, 1989).

Além disso, retomamos o pensamento de Guimarães (2017) acerca da dimensão cultural localizada nas questões sociais, políticas, éticas educativas e contextuais da organização do conhecimento, destacando o papel da organização do conhecimento no tocante do que possam integrar o plano de suas discussões e aplicabilidades. O que de certa forma, solicita a emergência de explorar os conceitos e as implicações acerca do papel da representação e organização do conhecimento no contexto étnico-racial.

Isso posto, apresentamos a seguinte questão: **De que maneira as pesquisas nacionais e internacionais em Análise de Domínio e construção de Tesauros contribuem e/ou divergem conceitual, teórico e metodologicamente das questões étnicas-raciais e culturais levantadas a partir das influências étnicas no léxico do português falado no Brasil?**

1.2 Justificativa

Diante da literatura, o presente trabalho está justificado pelas necessidades práticas relativas à elaboração de tesauros que utilizam ou visam utilizar os termos advindo do léxico do português influenciado pelas línguas dos povos originários e integrantes do Brasil. Em

realidade, busca-se reintroduzir aspectos conceituais, teóricos e metodológicos, com vistas à formulação de propostas atualizadas e condizentes com a realidade das instituições (bibliotecas, museus, arquivos e outros organismos culturais) e demais setores e organizações públicas e privadas.

Almejou contribuir com os trabalhos já publicados no tocante a construção de Tesouros como no caso do Modelo Metodológico Integrado para Construção de Tesouro proposto por Cervantes (2009), acrescentando outras perspectivas e elencando possíveis problemas e soluções para esse sistema de organização do conhecimento. Nessa perspectiva, nos motivamos pela busca, fundamentação e aplicações da Análise de Domínio na Ciência da Informação, mais precisamente nos estudos de sistemas de organização do conhecimento, posto considerarmos a possibilidade de inovações na sistematização de métodos e teorias para a construção de tesouros no cenário apresentado anteriormente.

Justificamos, ainda nossa proposta, na apresentação de novos tópicos e possíveis descobertas que promovem a acessibilidade aos sistemas de recuperação da informação, estimulando melhorias na organização e representação no âmbito da diversidade étnico-racial brasileira.

Ademais, ressaltamos o nosso interesse nessa temática, por acreditarmos que as reflexões oriundas desta pesquisa, poderão contribuir com a formação de profissionais e estudantes, somando as iniciativas destinadas às metodologias de construção de tesouros.

1.3 Hipóteses

A elaboração de sistemas de organização do conhecimento contempla a investigação e a gestão das estruturas conceitual e terminológica viabilizam a representação, organização e recuperação do conhecimento no contexto técnico-científico. Entretanto, a pesquisa levanta as seguintes **hipóteses**:

Hipótese 1 - A fundamentação e os aspectos metodológicos aplicados à elaboração de Tesouros em conjunto com o aporte teórico-conceitual e metodológico da Análise de Domínio se complementam e dialogam com a prática representacional e organizacional em domínios científicos e não científicos;

Hipótese 2 - Quanto a construção de tesouros que utilizam termos oriundos do léxico de influência étnico-racial e as investigações da Análise de Domínio na

Ciência da Informação, os aportes teóricos e metodológicos se apresentam insuficientes para tratar e investigar tal realidade étnico-racial no Brasil.

Assim, esta pesquisa defende a **tese** de que a Análise de Domínio aporta elementos necessários para a reestruturação teórico e metodológica dos tesauros no contexto do léxico do português brasileiro estabelecido pelas heranças linguísticas dos povos indígenas e afro-brasileiros.

1.4 Objetivos

Objetivo geral: analisar as possíveis contribuições, divergências e/ou aproximações teóricas, conceituais e metodológicas entre a construção de tesauros, a Análise de Domínio e os estudos dos léxicos a partir das influências étnicas brasileiras, considerando as pesquisas em âmbito nacional e internacional.

Os objetivos específicos que pretendemos alcançar são:

- a) Examinar aspectos teóricos, conceituais e metodológicos basilares na construção do tesouro de recuperação amplamente difundido na Organização do Conhecimento e nos padrões internacionais vigentes;
- b) Verificar, na literatura científica da Ciência da Informação tanto no contexto brasileiro, quanto internacional, a presença dos aspectos teóricos e conceituais da Análise de Domínio direcionados à construção de tesauros;
- c) Investigar os estudos e as frentes epistemológicas que discutem e contribuem com os usos e o entendimento dos léxicos influenciados pelos aspectos étnicos;
- d) Descrever as interseções e/ou inconsistências na construção de tesauros, a Análise de Domínio e os estudos dos léxicos diante das influências étnicas.

1.5 Estrutura da tese

Esta tese apresenta no terceiro capítulo a trajetória dos primeiros tesauros, percorrendo os primeiros inícios históricos de sua estruturação até o presente cenário de desenvolvimento dos vocabulários controlados e sistemas estruturados e conectados à internet. Nesse relato, são descritos os aspectos evolutivos, a fundamentação teórica e os padrões utilizados na construção de tesauros.

O quarto capítulo retoma a Análise de Domínio na Ciência da Informação conforme as discussões epistemológicas e a necessidade de reposicionar às pesquisas para o caminho dos domínios do conhecimento. Além disso, conceitos fundamentais e dimensões permitem elucidar como essa perspectiva é visualizada e moldada.

Adiante, o quinto capítulo trata dos aspectos e das questões acerca dos léxicos e das influências étnicas. Conceitos são descritos de modo a compreender o cerne central da temática étnico-racial e cultural, das quais implicam nas origens étnicas presentes no português falado no Brasil.

Os demais capítulos estão dedicados à pesquisa empírica, apontando os resultados, as análises e discussão dos dados, a síntese e a conclusão. As categorias principais (conceituais, teóricas e metodológicas) orientam o reconhecimento dos achados e a elucidação das inferências.

2 PROCEDIMENTOS METODOLÓGICOS

De início, entende-se que a pesquisa se molda a partir de projetos e ações resultando em um exercício intelectual. A pesquisa científica é concebida como um conjunto de processos sistemáticos e empíricos destinados ao estudo de um fenômeno (Sampieri; Collado; Lucio, 2013), refletindo assim em uma atividade essencial para o desenvolvimento científico.

Para esclarecer o percurso metodológico delineado aqui, Minayo (2009) entende que a metodologia é um caminho do pensamento e a prática exercida na abordagem de uma realidade, ao consistir, de forma simultânea, em teoria de abordagem (o método), de instrumentos para operar o conhecimento (as técnicas) e do exercício criativo do pesquisador (sua experiência e capacidade). O conjunto formado pelo método, as técnicas e a experiência, guia todos os passos necessários para prosseguir, desenvolver e relatar o conhecimento e a experiência captados pelo pesquisador, juntamente com o alcance dos seus objetivos e das respostas para as questões inicialmente levantadas.

A pesquisa em *lide*, caracteriza-se, segundo os objetivos, como uma pesquisa correlacional e, segundo as fontes, como uma pesquisa documental. A pesquisa correlacional refere-se ao “[...] tipo de estudo que tem como finalidade conhecer a relação ou o grau de associação existente entre dois ou mais conceitos, categorias ou variáveis em um contexto específico” (Sampieri; Collado; Lucio, 2013, p. 103). Entende-se que o estudo correlacional tem a possibilidade de mediar uma avaliação para perceber o tipo de vínculo, seja positivo ou negativo, entre o suposto relacionamento das partes (conceitos, categorias ou variáveis).

Desse modo, a pesquisa correlacional corresponde ao objetivo geral traçado para observar a relação entre a Análise de Domínio, os fundamentos de construção de Tesouros e os estudos léxicos que integram as influências étnicas. Essa pesquisa, também, possui a amplitude de estudar o uso da Análise de Domínio referente à abordagem de Classificações Especiais e Tesouros, mencionada por Hjørland (2002a) como uma das 11 abordagens utilizadas na Ciência da Informação e na Biblioteconomia.

Quanto à pesquisa documental, essa buscou investigar a literatura científica no tocante aos aspectos da construção de tesouros, da Análise de Domínio e dos estudos dos léxicos étnicos. Além disso, permitiu expor e examinar as conexões estabelecidas entre os conceitos, teorias e outras categorias importantes extraídas do material. A pesquisa documental atenta para fontes diversificadas e dispersas, considerando materiais que não recebem ainda um trabalho analítico e podem ser revisados de acordo com os objetos da pesquisa (Gil, 2002).

Para o referencial teórico, tomando como orientação a base do estudo de acordo com a revisão de literatura, buscou-se nas bases de dados, nos repositórios e bibliotecas digitais e universitárias, os termos alinhados à questão, aos objetivos e ao referencial teórico quais sejam: Análise de Domínios, construção de tesouros, os conceitos envolvendo a temática étnica e os léxicos étnicos.

Foi utilizada, inicialmente, para a etapa de elaboração do referencial teórico, as seguintes bases: Banco de Teses e Dissertações do IBICT, Web of Science, Scopus, Brapci (Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação), REDALYC - Red de Revistas Científicas de América Latina y el Caribe, España y Portugal, Internet Archive, ERIC (Education Resources Information Center) e *Google Scholar*. Além disso, a investigação buscou fontes documentais em bibliotecas e repositórios das universidades de ensino superior do Brasil, além de fontes informacionais como as bibliotecas e coleções digitais.

A pesquisa empírica foi realizada nas seguintes bases de dados: Base de Dados em Ciência da Informação (BRAPCI); Web of Science – Coleção Scielo Citation Index; Library, Information Science & Technology Abstracts (LISTA), aplicando estratégias booleanas (AND, OR, NOT) durante as buscas combinadas dos descritores (Tabela 1). Essas bases foram escolhidas em virtude de ambas apresentarem cobertura dos assuntos em diferentes continentes; tabulação de dados e metadados que facilitam a visualização; e a gestão da literatura mapeada. Dentre as facilidades dessas, os formatos dos arquivos de referências completas foram compatíveis e adaptados para edição em softwares e aplicativos bibliográficos.

Para a seleção dos documentos foram utilizados os seguintes critérios de inclusão: artigo, livros, resumos, capítulos, relatórios, normas, entre outros documentos; e referentes à literatura da Ciência da Informação e da Biblioteconomia. O material excluído correspondeu aquelas publicações pagas que não foram viabilizadas via acesso pelo Portal de Periódicos da CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) e ou publicações que não atentaram para o tema em questão.

Diante do material pré-selecionado, as orientações e etapas da análise de conteúdo pautaram a metodologia, além do mapeamento para viabilizar a organização das categorias de análise dos resultados e, conseqüente, a apresentação dos dados e das mensagens, as inferências e as discussões.

Para a coleta dos dados, elaborou-se uma ficha de registro documental que serviu como instrumento (roteiro) de apoio para a organização dos dados e mensagens. Esse roteiro forneceu ainda suporte para a estruturação das categorias analíticas, basilares e imprescindíveis para a análise de conteúdo. O detalhamento deste roteiro (Apêndice A) apresenta, além das

informações gerais dos documentos (título, autoria, ano de publicações), eixos responsáveis pelo agrupamento das categorias e dados registrados:

- a) Eixo I – Aspectos conceituais – 1.1 Conceitos, 1.2 Definições dos conceitos, 1.3 Pontos conflitantes, 1.4 Proposta;
- b) Eixo II – Aspectos teóricos – 2.1 Teoria, 2.2 Paradigma, 2.3 Abordagem teórica, 2.4 Epistemologias, 2.5 Pontos conflitantes, 2.6 Proposta;
- c) Eixo III – Aspectos metodológicos – 3.1 Abordagem metodológica, 3.2 Métodos, 3.3 Ferramentas ou instrumentos, 3.4 Padrões, 3.5 Diretrizes, 3.6 Pontos conflitantes, 3.7 Proposta.

Os eixos foram propostos de acordo com características gerais que cada conceito foi apresentado no referencial teórico. Assim, os aspectos gerais dos tesouros implicaram, por exemplo, na existência das diretrizes, padrões ou normas, características diferentes daquelas apresentadas durante o percurso introdutório da Análise de Domínio. Apesar dos eixos não incluírem tópicos iguais, ambos contêm elementos descritivos que possibilitaram aplicação desta pesquisa comparativa.

Apesar da coleta dados ser submetida a organização e registro nesses tópicos, não foi obrigatório o preenchimento de todos os tópicos para cada documento analisado, já que alguns trabalhos são exclusivamente teóricos, ou exclusivamente metodológicos, ou exclusivamente conceituais, outros documentos relataram todos os aspectos.

O estudo foi direcionado tanto para dados qualitativos, quanto quantitativos, pois ambos ampliam a apresentação e a consistência de variáveis e características das mensagens coletadas. O enfoque quantitativo é uma abordagem estruturada, sequencial e comprobatória representada por um conjunto de processos, visando a medição numérica e a análise estatística dos dados para estabelecer padrões e comprovar teorias (Sampieri; Collado; Lucio, 2013). Enquanto a qualitativa, procura tratar dados e mensagens conforme a dinâmica das relações e representatividade do objeto de pesquisa e dos fenômenos. Em outras palavras “[...] trabalha com o universo dos significados, dos motivos, das aspirações, das crenças, dos valores e das atitudes” (Minayo, 2009, p. 21).

Contudo, os dois tipos de enfoques podem ser utilizados em conjunto e permitiu um tratamento amplo e exaustivo, correspondente a multiplicidade de variáveis e categorias extraídas de diferentes tipos de dados e mensagens. Nessa perspectiva, concordou-se com a complementaridade destacada por Brufem (2001, p. 54):

Quanto ao papel dos valores na pesquisa, isto é, do ponto de vista axiológico, a complementaridade quantitativo-qualitativa transcende o maniqueísmo entre, por um lado, uma atitude totalmente liberta de valores como critério para obtenção da verdade e, por outro, uma situação de busca do conhecimento subjugada por escolhas ideológicas ou valorativas.

Apoiada na citação, acreditou-se que o enfoque misto assumiu aqui uma difícil tarefa para os pesquisadores. Tarefa essa, que solicitou, inicialmente, avaliar e questionar as possíveis formas e instrumentos no exercício empírico, buscando evidências relacionadas ao objeto de estudo.

Para auxiliar na coleta, tratamento e interpretação dos dados, a pesquisa retomou como orientação a Análise de Conteúdo na perspectiva de Bardin (2011). Ela surgiu no início do século XX diante de um cenário marcado pelo desenvolvimento do Behaviorismo na Psicologia e a disputa entre aqueles que defendiam a perspectiva quantitativa em oposição aqueles que afirmam a exclusividade e a necessidade de introdução da perspectiva qualitativa (Gomes, 2009). Essa análise tem origem da necessidade de tratar a experiência empírica diante do intenso fluxo das produções e a expansão dos meios de comunicação. Nesse intuito, buscou auxiliar psicoterapeutas, sociólogos, historiadores e outros pesquisadores das ciências sociais e humanas (Bardin, 2011).

Bardin (2011, p.48) define a análise de conteúdo como “um conjunto de técnicas de análise das comunicações visando obter por procedimentos sistemáticos e objetivos de descrição do conteúdo das mensagens indicadores (quantitativos ou não) que permitam a inferência de conhecimentos relativos às condições de produção/recepção (variáveis inferidas) dessas mensagens”. Por sua vez, Gomes (2009) explica ser uma técnica utilizada para analisar conteúdos de mensagens e uma possibilidade de compreender conteúdos na perspectiva qualitativa, além do auxílio de inferências que buscam explicar a dimensão da mensagem analisada. As principais maneiras de aplicação dessa técnica são: análise de avaliação ou análise de representacional; análise de expressão; análise de enunciação; e análise temática.

Hjørland (2002a) e Smiraglia (2015a) apontam que a análise de conteúdo se destacou com outras opções metodológicas e empíricas - dentre elas a análise de discurso e a teoria crítica, no auxílio do tratamento, leitura e procedimentos da Análise de Domínio.

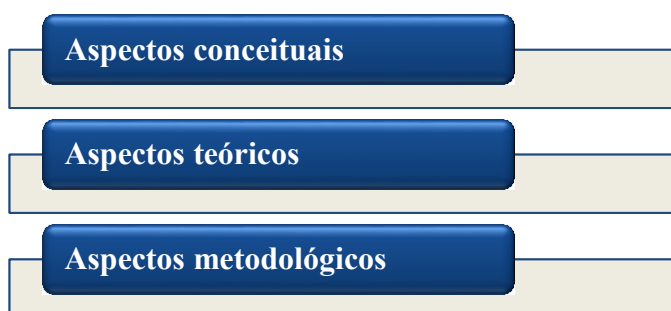
A atenção dada para as comunicações e a adaptabilidade diante das relações sociais e dos fenômenos proporcionam apreensão de conteúdos, levando em conta características que se modificam e permeiam entre os objetos e fenômenos estudados. A análise ainda busca introduzir um modo sistemático e exaustivo, considerando que a informação e os dados

somados detêm um impacto decisivo nas descobertas e no entendimento de questões específicas. A sistemática e a eficiência podem ser vistas na aplicação de regras regidas por concepções:

- homogêneas: poder-se-ia dizer que "não se mistura alhos com bugalhos";
- exaustivas: esgotar a totalidade do "texto";
- exclusivas: um mesmo elemento do conteúdo não pode ser classificado aleatoriamente em duas categorias diferentes;
- objetivas: codificadores diferentes devem chegar a resultados iguais; adequadas ou pertinentes: isto é, adaptadas ao conteúdo e ao objetivo (Bardin, 2011, p. 42).

É importante destacar essas regras são consolidadas nas etapas do trabalho e do exercício intelectual do pesquisador a partir de mensagens extraídas da literatura do caso das pesquisas bibliográficas, pois, nesta pesquisa, buscou-se identificar unidades importantes para investigar teorias, conceitos, descobertas e discussões relevantes durante o estudo empírico. As seguintes categorias foram utilizadas para agrupamento de dados, trechos de discurso, definições entre outras nuances utilizadas para embasar as inferências e interpretações:

Figura 1 – Categorias do estudo empírico



Fonte: Elaborada pela autora.

Diante das categorias da Figura 1, aplicou-se as técnicas de análises de conteúdo seguindo as etapas de: pré-análise da literatura; exploração dos documentos; tratamento dos dados – a essa etapa está relacionada a identificação e agrupamentos dos dados e mensagens em unidades (categorias e subcategorias analíticas); inferência e a interpretação (Bardin, 2011). Essas etapas atuaram em sequência e conforme as orientações precisas e bem definidas em cada uma delas, a considerar ainda a obediência as demais definições metodológicas.

Conforme a análise e a discussão dos dados, na categoria “Aspectos metodológicos”, ocorreu a necessidade de exemplificar métodos, análises e as demais ferramentas metodológicas por acreditar que através delas o entendimento e a perspectiva prática fossem descritos e explicados de forma clara e objetiva. Isso já reconhecendo métodos

e análises utilizadas na construção de tesouros, como já mencionadas na introdução aquelas relativas aos Estudos Bibliométricos, em específicos, análises e visualização das variáveis e dados por meio das redes de citação de autores. Os exercícios apontaram exemplos elucidativos das etapas, matrizes de cruzamento dos dados e frequências. Buscou-se, ainda, introduzir o léxico do português e algumas traduções de outros léxicos.

Utilizou-se ferramentas e softwares: VOSviewer (software gratuito para visualização e elaboração de redes bibliométricas), XMind (software com licenças pagas e gratuitas para editar e gerenciar de mapas conceituais e mentais), e o Mendeley (gerenciador gratuito de referências).

2.1 Pesquisa empírica

Conforme a aplicação das estratégias booleanas, as pesquisas nas bases de dados apresentaram resultados correspondentes às categorias conceituais desta pesquisa. As estratégias não limitaram o período, sendo abrangentes para todos os anos de publicações cobertos pelas bases. A Tabela 1 ilustra o quantitativo de documentos recuperados e os candidatos a serem analisados nas etapas das Análises de Conteúdo:

Tabela 1 – Estratégias booleanas e resultados

Temáticas analisadas	Base	Estratégia Booleana	Nº de resultados	Nº de pré-selecionados
Análise de Domínio	LISTA	(theory OR epistemology OR approach OR method OR methodology) AND (domain analysis) (todos os campos)	205	40
	Web of Science (Scielo Citation Index)	(TS= "domain analysis") AND (TS=theory) OR (TS= "domain analysis") (TS= epistemology) OR (TS= "domain analysis") (TS=approach) OR (TS= "domain analysis") (TS=method) OR (TS= "domain analysis") (TS= methodology)	138	77
	Brapci	"análise de domínio" AND teori* "análise de domínio" AND epistemolo* "análise de domínio" AND abordage* "análise de domínio" AND meto* (todos os campos)	77	41
Tesouros	LISTA	(theory OR epistemology OR approach OR method OR methodology OR construction OR standards OR guidelines) AND (thesaurus OR thesauri) (todos os campos)	682	80
	Web of Science (Scielo Citation Index)	(TS= thesaur*) AND (TS=theory) OR (TS= thesaur*) (TS= epistemology) OR (TS= thesaur*) (TS=approach) OR (TS= thesaur*) (TS=method) OR (TS=thesaur*) (TS= methodology) OR (TS= thesaur*) AND (TS=construction)OR (TS= thesaur*) AND (TS=standars) OR (TS= thesaur*) AND (TS=guidelines)	545	95

	Brapi	"tesaur*" AND teori* "tesaur*" AND epistemolo* "tesaur*" AND abordage* "tesaur*" AND meto* "tesaur*" AND padro* "tesaur*" AND diretri* (todos os campos)	172	69
Temáticas analizadas	Base	Estratégia Booleana	Nº de resultados	Nº de pré-selecionados
Etnia, Raça, Cultura e Análise de Domínio	LISTA	(ethnicity OR ethnicities OR race OR culture) AND (domain analysis) (todos os campos)	21	10
	Web of Science (Scielo Citation Index)	(TS=("domain analysis")) AND TS=(culture)OR (TS=("domain analysis")) AND TS=(race) OR (TS=("domain analysis")) AND TS=(ethni*)	12	7
	Brapi	"análise de domínio" AND etni* "análise de domínio" AND raça "análise de domínio" AND cultur*	14	9
Etnia, Raça, Cultura e Tesauros	LISTA	(ethnicity OR ethnicities OR race OR culture) AND (thesaurus OR thesauri) (todos os campos)	54	8
	Web of Science (Scielo Citation Index)	(TS=(thesaur*)) AND TS=(culture)OR (TS=(thesaur*)) AND TS=(race) OR (TS=(thesaur*)) AND TS=(ethni*)	17	9
	Brapi	"tesaur*" AND etni* "tesaur*" AND raça "tesaur*" AND cultur*	18	10
Análise de Domínio e Tesauros	LISTA	(domain analysis AND thesaurus) OR (domain analysis AND thesauri) (todos os campos)	17	6
	Web of Science (Scielo Citation Index)	(TS=(domain analysis)) AND TS=("thesaur*")	9	4
	BRAPCI	"análise de domínio" AND "tesaur*"	3	
Análise de Domínio, Tesouro e Etnia	LISTA	(domain analysis) AND (thesaurus) AND (ethnicity)	0	0
	Web of Science (Scielo Citation Index)	((TS=(domain analysis)) AND TS=(thesaur*)) AND TS=(ethni*)	0	0
	BRAPCI	"análise de domínio" AND "etni*" AND "tesaur*"	1	1
Léxico	LISTA	(theory OR epistemology OR approach OR method OR methodology OR construction OR standards) AND (lexicon)	284	37
	Web of Science (Scielo Citation Index)	(TS= lexicon) AND (TS=theory) OR (TS= lexicon) (TS= epistemology) OR (TS= lexicon) (TS=approach) OR (TS= lexicon) (TS=method) OR (TS=lexicon) (TS= methodology) OR (TS=lexicon) AND (TS=standars) OR (TS=lexicon) AND (TS=guidelines)	305	24
	BRAPCI	"léxico" AND teori* "léxico" AND epistemolo* "léxico" AND abordage* "léxico" AND meto* "léxico" AND padro* "léxico" AND diretri* (todos os campos)	10	5
Léxico, Análise de domínio e Tesauros	LISTA	(domain analysis) AND (thesaurus) AND (lexicon)	2	0
	Web of Science (Scielo Citation Index)	((TS=(domain analysis)) AND TS=(thesaur*)) AND TS=(lexicon)	1	1
	BRAPCI	"análise de domínio" AND "léxico" AND "tesaur*"	0	0

Temáticas analisadas	Base	Estratégia Booleana	Nº de resultados	Nº de pré-selecionados
Léxico, Etnia, Raça e Cultura	LISTA	(ethnicity OR ethnicities OR race OR culture) AND lexicon	27	0
	Web of Science (Scielo Citation Index)	(TS=(lexicon)) AND TS=(ethni*) OR (TS=(lexicon)) AND TS=(culture) OR (TS=(lexicon)) AND TS=(race)	61	12
	BRAPCI	léxico AND raça léxico AND etni* léxico AND cultura	2	0
Léxico, Análise de Domínio, Tesouros e Etnia	LISTA	(domain analysis) AND (thesaurus) AND (lexicon) AND	0	0
	Web of Science (Scielo Citation Index)	((TS=(domain analysis)) AND TS=(thesaur*)) AND TS=(lexicon)) AND TS=(ethni*)	0	0
	BRAPCI	"análise de domínio" AND "léxico" AND "tesaur*" AND "etn*"	0	0
Total			2677	545

Fonte: Elaborado pela autora.

A utilização das dez temáticas citadas justifica-se por compreender os principais termos mencionados no referencial teórico² e a necessidade de verificar alguns assuntos separadamente, como foi o caso de documentos na CI e na Biblioteconomia que trataram somente dos léxicos, e de modo combinado para estabelecer interseções e diálogos. Essa estratégia geral aplicada as temáticas serviram para mensurar a quantidade de publicações seja isoladamente ou absorvendo outra temática nacional ou internacional. Podemos visualizar que a “Léxico + Análise de Domínio + Tesouros + Etnia” não recuperou nenhum trabalho. Enquanto, a temática “Etnia + Raça + Cultura + Tesouros” aparece em poucas publicações e pode ser vista como um assunto ainda em desenvolvimento.

Foram recuperados no total de 2677. Após a leitura flutuante, 545 documentos foram pré-selecionados por corresponderem aos critérios de: serem artigos, livros, resumos, capítulos, relatórios, normas ou outro tipo de documento relevante para a temática; e publicações referentes à literatura da Ciência da Informação e da Biblioteconomia. Os conceitos e sintagmas destacados no referencial foram utilizados como subsídio a serem verificadas nos documentos pré-recuperados, dentre eles:

² A indicação dos termos das temáticas seguiu a frequência com que eram apresentados na literatura durante a fase de exploração, mesmo considerando pontos negativos em relação a determinados termos, como o caso de “raça”. Percebeu-se que alguns termos eram mais incidentes do que outros, embora o termo raça, também, tenha recuperado documentos e outros termos como negritude e afro-brasilidade.

- a) Tesouros - termo, conceito, sintagmas, categoria, faceta, eixos e relações paradigmáticas e sintagmáticas, entidades, definição, padrões, normas, diretrizes, métodos, ferramentas, etapas de construção de tesouros;
- b) Análise de Domínio - domínio, comunidade discursiva, comunidade epistêmica, gênero textual e documental, dimensão sociológica, dimensão ontológica, dimensão epistemológica;
- c) Etnicidade e léxicos - etnia, étnico, etnicidade, raça, racismo, indígenas, afro-brasilidade, negritude, classificação social, vocabulário, léxico, línguas naturais, variantes lexicais, etnoconhecimento.

Essas unidades foram buscadas durante a leitura flutuante e o preenchimento da Ficha de registro documental (Apêndice A). Após esse procedimento, verificou-se que somente 139 documentos, em sua maioria artigos e capítulos de livros, corresponderam e atenderam aos elementos investigados. A análise prosseguiu quando pelo menos dois autores, em pelo menos uma temática, apontou e desenvolveu o conceito, a teoria ou outra indicação conceitual, teórica ou metodológica.

3 ASPECTOS EVOLUTIVOS E DE CONSTRUÇÃO DE TESAuros

Os primeiros indícios do uso do termo “*thesaurus*” e a ideia por trás de sua atual proposta têm origens nas publicações que objetivavam apresentar e sistematizar um determinado conhecimento de uma língua. O termo *thesaurus* tem origem do latim e remete-se ao *thesaurós* do grego e, em alguns dicionários ao *tesoro*, foi esse último termo que o poeta e chanceler florentino Brunetto Latini (1220-1294) nomeou a obra enciclopédica “*Livre dou Trésor*” escrita em francês entre os anos de 1262 a 1268 e, posteriormente, na Espanha o jurista Vidal de Canvelles utilizou o seu equivalente em latim na obra “*Compilatio Maior o In excelis Dei Thesaurus*” publicada no século XIII (Martínez de Sousa, 1989).

Thierry e Robert Etienne publicaram em 1532 o *Thesaurus Latinde Linguae*, dicionário em língua latina era estruturado em arranjo alfabético. Contudo em 1572, Henri Etienne II entusiasmado com o trabalho iniciado por seu pai Robert, continua a proposta e publica o *Thesaurus language Graecae* (Motta, 1987). O reflexo desse tipo de publicação pode ser compreendido com o avanço da elaboração de dicionários e enciclopédias acompanhadas pelas tecnologias de impressão e os movimentos científicos da época que intensificaram com o advento do Iluminismo.

Referências recordam outras formas e propósitos que reuniam palavras da linguagem natural. Dicionários, enciclopédias, glossários, índices, classificações, objetos do exercício da Lexicografia e da Terminografia³ e outros ramos e disciplinas, ilustram bem esse conjunto de referências do qual antecede a evolução dos tesauros de recuperação. Embora a centralidade dos estudos atuais enfatiza seu foco na indexação, ou seja, na representação temática e suas consequências para a classificação, categorização e acesso de diferentes tipos de documentos, vestígios podem ser mapeados em decorrência da proximidade com as investigações dos conceitos no entendimento próximo ou acerca dos estudos léxicos, antes mesmo das análises fixadas nas bases conceituais e científicas relatadas com maior ênfase na Terminologia do século XX.

A classificação temática dos vocabulários com o intuito do agrupamento e relacionamento de palavras remonta ao dicionário de Sânscrito Amara Cosha, ao *Onomastikon* de Pollux e ao glossário latino anglo-saxão de Elfrico de Eynsham (Jones, 1972). Em

³ Esclarecendo, a Lexicografia é um ramo aplicado da Lexicologia responsável pela elaboração de dicionários, já a Terminografia é o ramo aplicado da Terminologia responsável pela elaboração de dicionários e glossários terminológicos em contextos especializados. (CABRÉ, 2005).

Onomastikon por exemplo, Pollux elaborou uma obra lexicográfica sobre a cultura grega na forma de 10 grandes classes (10 volumes) estabelecidas para agrupar palavras em assuntos (König, 2016), ao invés da disposição em ordem alfabética e do desprezo das relações entre palavras com significados semelhantes. Isso remete à organização de vocabulários para diferentes fins e com diferentes funções ao longo da história. Demarca também, que parte integrante desses registros documentaram o uso de palavras de determinadas línguas, contribuindo para a consolidação de idiomas presentes atualmente. Em reflexão, o uso de um conjunto de vocabulário em um contexto linguístico definido indica heranças para estudos recentes, visto que parte do caráter informativo e conceitual permite a distinção e a identificação de grupos de palavras.

Há um entrelaçamento de trabalhos com o foco nos agrupamentos de palavras, sejam eles explícitos ou implícitos, dos quais influenciaram a trajetória de desenvolvimento dos tesouros. A primeira tendência identificada por Jones (1972) é a sinonímia, essa como característica da linguagem natural que viabiliza relacionamentos verbais pela proximidade de palavras com o mesmo sentido e semelhança na base geral de um léxico. Posteriormente, os dicionários ganharam notoriedade na identificação, ainda que de forma indireta, das relações verbais. A terceira, o desejo geral em tomar como base argumentos sólidos e explícitos descritos na escrita filosófica, sentindo a necessidade de o discurso científico ser capaz de lidar nos modos de descrição, além da influência do caráter universal e o encontro com as origens conceituais e a base semântica (Jones, 1972).

O ‘*Thesaurus of English Words and Phrases*’ desenvolvido por Peter Mark Roget em 1852 configura-se como um dos mais famosos e é o mais mencionado na literatura. A ideia de Roget era apresentar de modo sistemático palavras que melhor transmitisse uma determinada ideia (Aitchison; Clarke, 2004). Sua principal função era fornecer um instrumento de consulta e facilitar a escolha de palavras adequadas ao contexto de leitores e pesquisadores, ou seja, uma obra de referência para a literatura inglesa.

Entretanto, o tesouro literário, como o de Roget, apesar de relacionar e recuperar sinônimos, difere em termos de finalidade, pois na recuperação o que interessa são os termos e os vocabulários elegidos para entradas e para a indexação em situações restritas e especializadas (Jones, 1972). Vickery (1960) observou que essa obra possui as seguintes características: a ligação das palavras-ideias com palavras-texto; as palavras-ideias são classificadas; e por último a estruturação de um índice alfabético para todas as palavras, indicando assim uma rede de conexões que possibilitou algumas reflexões no trabalho de pesquisadores como o Peter Luhn.

A proposta de Roget realmente marca um tipo de tesouro geral que viria a amadurecer a idealização dos tesouros de recuperação durante o século XX. Além disso, Jones (1972, p. 400, tradução nossa) esclarece que o tesouro de recuperação “[...] distinto de uma lista de cabeçalhos de assunto ou classificação completa, é uma introdução relativamente recente, embora seus parentes de indexação remontam ao século XIX, cabeçalhos de assunto a Cutter em 1876 e classificações a Dewey em 1887”. Cerca de um século depois o termo foi aplicado ao contexto das listas de vocabulários utilizados na recuperação da informação (Aitchison; Clarke, 2004).

Segundo Roberts (1984), Lancaster (1986), Aitchison e Clarke (2004), Clarke (2019) o tesouro de recuperação da informação moderno tem origens na década de 1950, período marcado pela necessidade de transformação e adoção de sistemas de recuperação mais aprimorados ao crescimento documental conforme o contexto técnico e científico esboçado durante o Pós Segunda Guerra Mundial.

Fatos ocorridos durante as primeiras décadas do século XX promoveram mudanças nos setores técnico científico, sociopolítico e econômico, sobretudo, em decorrência do avanço industrial e militar consequentes das duas grandes guerras mundiais e do fortalecimento de instituições e do sistema econômico vigente. O que levaria a uma corrida por garantir eficácia das tecnologias de comunicação e processamento de dados e informações, além do desenvolvimento dos sistemas de recuperação correspondentes ao momento de criação e expansão dos computadores.

Em meados de 1950, as exigências direcionadas ao desenvolvimento dos sistemas de recuperação trouxeram o abandono de princípios ortodoxos de recuperação em muitos ambientes especializados e deram ênfase ao controle terminológico e à estruturação conceitual como novas ortodoxias incorporadas pelo tesouro de recuperação da informação (Roberts, 1984). A lista de cabeçalho de assunto e o esquema de classificação são ferramentas tradicionais de recuperação com propósitos e formas estruturais próprias, o cabeçalho fornece uma lista de palavras-chaves ordenadas alfabeticamente e relacionadas de acordo com os seus significados, enquanto os esquemas se qualificam como listas sistemáticas utilizadas na organização em classes e uso de códigos (Vickery, 1960). Portanto, essas exigências deveriam superar o uso padrão dos assuntos de cabeçalhos e das classes estabelecidas como tipos de etiquetas e agrupamentos, permitindo, por outro lado, vincular respostas dos usuários à rede de termos e palavras utilizadas na indexação.

Roberts (1984) foi muito analítico e assertivo quando mencionou haver uma confusão terminológica acerca do termo ‘thesaurus’ empregado e descrito em diferentes

propostas. É natural que outros termos sejam denominados, como o caso de ‘dicionário de sinônimos’ ser um equivalente da tradução do inglês para o português e o uso de vocabulário controlado como equivalente terminológico. Podemos encontrar uma vasta literatura mencionando e tomando como objeto o “thesaurus”, embora deva-se compreender se o conceito e o termo são aplicados para a finalidade e a função das quais a maioria dos autores compartilham na Ciência da Informação e Biblioteconomia. Há divergências e pontos distintos que desenham os alicerces e a fundamentação teórica, dentre aquelas alocadas na Linguística Documentária, Teoria Linguística e Léxico Documentário (Gardin, 1966; 1973; García Gutiérrez, 1990) até as recentes investigações da *Knowledge Organization* (KO) (Moreira, 2019; Clarke, 2019).

Foskett (1996), Aitchison e Clarke (2004) atribuíram a Hans Peter Luhn como o primeiro a utilizar “*thesaurus*” no contexto moderno. Embora, Roberts (1984) sugere que o crédito por introduzir o tesouro na literatura profissional da época recai para as propostas apresentadas por Charles Bernier e Evan Crane, do Chemical Abstracts Service, em 1947 a partir da originalidade da aplicação no âmbito da recuperação. Ainda considerando que Calvin Moores tenha divulgado suas ideias em 1947. A ideia inicial era de projetá-lo para aplicação tanto na indexação, quanto na pesquisa, utilizando a operação de coordenadas, essa última orientada por Batten e Mooers em sistema pós-coordenado e manual (Aitchison; Clarke, 2004).

O desenvolvimento dos primeiros tesouros de recuperação de informação teve início no Centro de Informação de DuPont em 1959, seguido do Tesouro de Descritores da ASTIA (*Armed Services Technical Information Agency*) de 1960 e do Dicionário de sinônimos de Engenharia Química do *American Institute of Chemical Engineers* de 1961, esses tiveram mais reconhecimento como os primeiros tesouros de recuperação (Roberts, 1984; Aitchison; Clarke, 2004).

Um fato importante pode ser destacado: o distanciamento desses tesouros daqueles outros empregados somente para o uso da linguagem natural, pois no contexto moderno o que influenciava era a linguagem utilizada nos centros e organizações científicas e de natureza especializada e terminológica, não as funções literária e de consulta de sinônimos. Os grupos científicos e políticos, bem como associações, instituições comerciais (nos casos das empresas Du Pont e IBM - International Business Machines Corporation) e órgãos políticos, seriam os principais beneficiários dos tesouros.

E assim, nas décadas seguintes seriam amplamente introduzidos em bibliotecas públicas e para outros públicos. Broughton (2006a) explica que durante a década de 1960 começou-se a incorporar princípios estruturais e a dispor as unidades em referências cruzadas,

dos quais serviram para o formato do *Thesaurus of Engineering and Scientific Terms*, publicado em 1967, cujo diferencial foi a descrição de regras que serviram como base para sua criação.

Há uma ênfase na literatura conforme a leitura de Roberts (1984), Vickery (1960), Aitchison e Clarke (2004) e Foskett (1996) em apresentar o contexto dos Estados Unidos com território fértil e provedor do aparecimento dos tesouros, possivelmente, por haver uma reivindicação e divulgação forte entre as instituições e nos eventos norte-americanos. Apesar dos estudos britânicos também serem importantes nesse período de amadurecimento.

3.1 Principais vertentes

É inegável que o terreno fértil para o aparecimento dos primeiros tesouros de recuperação tenha sido na América do Norte, em específico nos Estados Unidos (Campos, 2001). Grande incentivo esteve concentrado no financiamento de pesquisas e no desenvolvimento de novas tecnologias de informação, além de evidenciar as pesquisas norte-americanas. Por consequência, a abertura dada para potencializar mecanismos de gestão e recuperação de documentos, aliados ao aperfeiçoamento dos computadores, às bases de dados e aos índices e resumos científicos, permitiram mudanças mais significativas. Isso possibilitaria deixar de lado o imperativo de outras tecnologias como os microfimes e os pontos de acessos em fichas analógicas, o que por outro lado, essas tecnologias seriam associadas e reformuladas em outros processos legíveis por máquinas computadorizadas.

Considerando o contexto da CI nos Estados Unidos e outros países por suas vertentes, Day (2001) revela que nesse período a Teoria da Matemática da Informação e a Cibernética, principais expoentes teóricos, despontavam no cenário e viriam a ser influências presentes na história dos discursos e das teorias da informação. Assim, as preocupações com os canais de comunicação e a necessidade de transmissão da informação entre os indivíduos podem ser considerados fatores que permearam nos tesouros estadunidenses.

Considerando os antecedentes, a catalogação de assuntos era o principal instrumento de organização utilizado nas bibliotecas e centros de documentação durante o final do século XIX e início do século XX. Gomes e Marinho (1984) esclarecem que esse sistema de catálogos de assunto surgiu enquanto um sistema pré-coordenado desenvolvido no século XIX na Library of Congress (Biblioteca do Congresso Americano), tendo como influência os ideais da Revolução Francesa em considerar as bibliotecas como instituições com funções sociais e destinadas a educação do povo, logo disponibilizar técnicas e serviços destinados ao acesso livre às estantes e os serviços de referência.

Havia uma série de contribuições à catalogação de assunto, dentre as quais: a utilização das remissivas, as entradas analíticas e a divisão de listas alfabéticas (Mey; Silveira, 2009). Muitos movimentos e congressos estiveram intensificando a catalogação de assunto e a catalogação com o foco na descrição física. Parte dessas iniciativas eram propostas por instituições e de grupos com a American Library Association⁴, fundada em 1876 por Cutter, Melvil Dewey e outros bibliotecários, e da Library of Congress.

Notou-se dificuldades que indicariam necessidades de padronização e de tratamento dos assuntos. As listas alfabéticas de cabeçalhos de assuntos e os esquemas classificatórios não priorizavam algumas características na linguagem natural. Nesse sentido, Vickery (1960) observa que esses ignoravam a dispersão de sinônimos e estavam articulados, essencialmente, em padrões de palavras-chave e pouco auxiliavam na operação de ligação entre palavra-texto e a palavra-chave atribuída pelo indexador.

Em 1951, Mortimer Taube elaborava o *Uniterm* e o introduzia na indexação por meio de unidades assuntos estabelecidos e ligados em referências cruzadas. Em sua publicação de 1952 intitulada *Unit Terms in Coordinate Indexing* apontou a ideia concentrando-se no que ele chamou de “indexação coordenada”, como o termo já indica seria uma forma de coordenação dos termos durante o tratamento indexical.

Taube estudou os sistemas de indexação para seu trabalho desenvolvido no Armed Forces Technical Information Agency (Centro de Informações Técnicas de Defesa dos Estados Unidos) e cria esse sistema baseado em termos enquanto unidades que viriam superar os problemas da ordem dos cabeçalhos de assuntos, aderindo à extração de palavras do texto e a associação dessas palavras em todo o sistema (Wright, 1972). O sistema Uniterm de Taube também apresentou dispersões léxicas e outras questões em virtude da extração de unidades simples, das quais considerando os aspectos léxicos e semânticos não funcionavam com eficiência.

Os sinônimos e outras dispersões léxicas configuram-se enquanto obstáculos na representação. Adaptações de técnicas e procedimentos as ferramentas tecnológicas da época, também, encaminharam a transição da catalogação de assunto e do sistema Uniterm para os projetos tesaurais (Campos; Gomes, 2006). Nessa perspectiva e seguindo o avanço das tecnologias de leitura e mecanização dos sistemas de recuperação, Luhn aponta uma primeira

⁴ A ALA foi fundada em 6 de outubro de 1876 durante um congresso na Filadélfia com a participação de bibliotecários e nomes fundadores como Justin Winsor (Boston Public, Harvard), William Frederick Poole (Chicago Public, Newberry), Charles Ammi Cutter (Boston Athenaeum), Melvil Dewey (Amherst) e Richard Rogers Bowker (Publisher's Weekly) (American Library Association, 2021).

sugestão de que um tesouro poderia permitir que uma máquina traduzisse palavras-texto para palavras-chave, viabilizando, posteriormente, sua proposta de estabelecimento de um dicionário capaz de agrupar palavras com significados semelhantes (Vickery, 1960).

No Zatocoding de Mooers e no dicionário de Luhn, os termos preferidos dispostos no índice receberam significados mais amplos e, também, associados a outros termos equivalentes de recuperação, além disso a precisão foi alcançada através da coordenação de termos amplos e na época não possuíam recursos estruturais mais sofisticados do que os seus sucessores (Roberts, 1984).

Além disso, a limitação do *Uniterm* daria espaço para a discussão voltada para os termos compostos e a redução do sentido quando somente termos simples (uma única palavra) seriam coletados para as listas de assuntos. Lancaster (1986) aponta problemas comuns nos sistemas que utilizavam uma única palavra, dentre eles: os homógrafos com mesma grafia e significados completamente diferentes; e os problemas em decorrência do ponto de vista do qual as palavras simples poderiam se adequar em muitos contextos. Muitos desses problemas já eram comuns para indexadores.

Para evitar tais impasses os tesouros foram introduzidos enquanto uma forma de orientação externa aos índices, sendo viáveis para o processamento dos computadores e a padronização de entradas (Bernier, 1968). O impacto não seria somente visto na disponibilização de índices para usuários e indexadores, mas na economia do trabalho de indexação, já que os termos e suas redes de relações estariam no sistema e ficariam disponíveis para o manuseio.

Bernier deu continuidade na ideia que antes havia esboçado e em julho de 1957 utilizou o termo '*technical thesaurus*' no vocabulário de recuperação e refletiu acerca de outros autores, anteriormente, terem negligenciado a natureza das relações entre os conceitos. Além de Bernier e Crane, anteriormente, terem defendido uma forma híbrida que buscava a introdução do conceito associativo com auxílio de um esquema convencional de relações conceituais preestabelecidas (Roberts, 1984).

Charles L. Bernier, Evan J. Crane, Mortimer Taube, Calvin N. Mooers e Hans Peter Luhn tinham algo em comum em seus trabalhos: ambos desenvolviam pesquisas e trabalhavam diretamente com indexação no contexto de áreas e especialidades. Crane⁵, por exemplo, teve

⁵ O American Chemical Society (2021) detalha a atuação de Crane e cita também Bernier com um dos seus parceiros de trabalho e, também, diretor do Chemical Abstracts Service.

um papel muito importante na área da Química através do seu trabalho na Chemical Abstracts Service e nas tendências de resumo e indexação para revistas científicas.

O contexto europeu, principalmente o britânico, também esteve mantendo o foco na construção de listas de assuntos que auxiliassem na tradução e síntese, contudo havia aspectos teóricos e metodológicos distintos da origem dos tesouros norte-americanos. O principal aspecto foi a introdução da classificação bibliográfica, apontada nas análises evolutivas de Lancaster (1986).

As mudanças ocorridas estavam centralizadas na remodelação de cabeçalhos de assuntos sistemáticos, diferentes daqueles organizados como uma lista alfabética e de modo linear. Segundo Campos (2001) o grupo Classification Research Group foi entusiasta dos suportes teóricos da época e concentrou suas atividades no desenvolvimento de esquemas especializados em classificação, nas pesquisas sobre sistemas nocionais para tabelas de classificação e nos estudos de análise de conceitos.

Membros do Classification Research Group, Eric J. Coates, Douglas John Foskett e Derek Austin foram nomes de destaque diante dos aportes que subsidiaram as tendências de análise e elaboração de estruturas focadas nos assuntos e na forma integrada e sistemática que eles deveriam compor os sistemas. Tendo em vista os termos compostos e os índices de cadeia, Coates além de introduzir uma perspectiva inglesa dos pontos de vista de Ranganathan, em específico da classificação e análise facetada, ele conseguiu ampliar essa teoria subjacente e contribuir com os trabalhos de Jones, Austin, Immroth e outros (Svenonius, 1990).

A partir daí a influência da visão de classificação e de facetas de Ranganathan aos poucos implicou em outras formas de se pensar a organização e a composição das listagens de assuntos, fazendo com que os tesouros elaborados, posteriormente, retomassem as expressões, entidades, nomes e objetos concretos e abstratos.

Whelan argumentou (1958 *apud* Roberts, 1984, p. 278, tradução nossa) “[...] que a classificação hierárquica falhou em atender aos requisitos especializados das circunstâncias modernas de informação em grande parte por causa de inflexibilidades inerentes. A alternativa era empregar um arranjo em rede de índice de termos por causa da maior liberdade de conexões conceituais implícitas em tal forma”. O que leva a pensar que durante as fases iniciais de utilização da concepção de hierarquia herdada a partir da Classificação, alguns empasses não foram de fato solucionados. Tais motivos têm facilitado a aplicabilidade da análise de faceta ao encontro das possibilidades estruturais relativas às conexões conceituais.

Embora não sendo um trabalho amplamente divulgado, Roberts (1984) aponta que o primeiro trabalho experimental com um tesouro de recuperação foi elaborado pelo centro de

pesquisa Royal Radar Establishment⁶ na cidade de Malvern em 1955. Já em meados de 1950, pesquisadores da *Cambridge Language Research Unit* (CLRU) experimentaram um tesauro ligado a um sistema de coordenadas de cartões perfurados, utilizando uma metodologia de extração de termos índices e termos preferidos e uma das principais preocupações do grupo era com os aspectos relacionais (Roberts, 1984).

Aspectos estruturais perpassaram por modificações de acordo com a reorganização da exibição dos termos e a introdução de princípios fixados no estabelecimento das relações conceituais. O trabalho de Jean Aitchison destaca-se a partir de seu trabalho na English Electric Company na década de 1960 com exibição alfabéticas e sistemáticas interdependentes e o desenvolvimento do Thesaurofacet inspirado nos princípios analíticos de facetas (Broughton, 2006a).

O Thesaurus Facet apresentava uma tabela exclusiva em classificação organizada por facetas e uma listagem tesaural em ordem alfabética. A perspectiva facetada foi um marco histórico no processo evolutivo dos tesauros, por buscar subsídio na Teoria da classificação Facetada de Ranganathan para refinar a disposição dos conceitos. Aitchison, posteriormente, aplicou sua experiência culminando no tesauro da UNESCO na década de 1970 (Lancaster, 1986).

Urdiciain (1997) apresenta a origem espanhola divulgada na América Latina. Nessa perspectiva, Lancaster (1986) aponta duas vertentes responsáveis por influenciar a evolução dos padrões e surgimento dos tesauros no contexto moderno: a primeira corresponde ao índice alfabético de assunto atribuído ao âmbito dos Estados Unidos, e a segunda referente à classificação bibliográfica difundida no território britânico.

3.2 Teorias fundamentais na construção de tesauros

O desenvolvimento das metodologias, princípios e padrões aplicados à construção dos tesauros aconteceu de acordo com o estabelecimento de teorias e abordagens conduzidas na Ciência da Informação e Biblioteconomia. Esse apoio teórico, conceitual e prático esteve presente na forma como os profissionais executavam as etapas e todas as ações encontradas desde o tratamento de vocabulários e terminologias até as conexões estabelecidas pelos

⁶ O Royal Radar Establishment correspondeu em 1953 ao Royal Radar Establishment resultante da união entre dois centros ingleses, o Telecommunications Research Establishment e o Radar Research and Development Establishment, passando a coordenação do Ministério da Aviação, Ministério da Tecnologia, Ministério do Abastecimento da Aviação e em 1971 para o Ministério da Defesa. Fonte: The National Archives (2023).

relacionamentos e os diferentes tipos de esquemas sistemáticos, facetados ou alfabéticos. A seguir apresentaremos três teorias fundamentais para a consolidação dos tesouros enquanto instrumentos terminológicos e de organização do conhecimento.

3.2.1 Teoria Geral da Terminologia

A Teoria Geral da Terminologia, ou Teoria Clássica da Terminologia, inaugurou as primeiras atividades da terminologia enquanto disciplina científica na modernidade e possibilitou algumas contribuições no campo da Ciência da Informação e Documentação. Surgiu em meados da década de 1930, em virtude de um contexto responsável por fomentar a comunicação científica imprescindível para acompanhar o progresso e a atuação de diferentes disciplinas e setores das ciências e tecnologias.

O engenheiro austríaco Eugen Wüster intrigado com as questões equívocas envolvendo os processos de comunicação reconhece a necessidade de estudos e mudanças significativas no sentido de priorizar a descrição e a padronização de termos. Nesse sentido, Wüster elabora os primeiros pressupostos de uma teoria da terminologia de acordo com sua experiência terminográfica ao compilar a obra *The Machine Tool: An interlingual Dictionary of Basic Concepts* e promover a Escola de Viena (Cabré; Azevedo, 2019). Nos anos seguintes o trabalho de Wüster ganhou notoriedade e consolida-se como a Teoria Geral da Terminologia (TGT) e o com o INFOTERM⁷ apoiado pela UNESCO.

Essa teoria recorre a ideia da comunicação inequívoca e universal por formalizar o trabalho de compilação, registro e tradução do qual facilitaria um processo normativo naquela época. O trabalho terminológico é assim considerado “[...] um elemento inerente às chamadas comunicações especializadas, as quais são tradicionalmente associadas à redação de artigos científicos, teses, resenhas, manuais, textos especializados em geral” (Krieger; Finatto, 2004, p. 16). A ligação fortemente estabelecida a partir dos registros das comunicações, principalmente artigos e teses, de certo modo apresentou-se como um elo com as teorias e a prática da organização e da representação da informação. Se o pensamento de Wüster era viabilizar os atos comunicativos no âmbito da especialidade, sejam eles escritos ou falados, através da compilação e padronização, os indícios tesauroais atestam o uso e o entendimento das

⁷O INFOTERM foi um órgão internacional responsável pelas questões de terminologia, envolvendo documentos, bancos de dados e vocabulário, além de também auxiliar nos projetos orientados para a análise dos conceitos (DAHLBERG, 1978c)

linguagens de especialidades também voltadas para esses atos. Essas linguagens funcionam como protocolos e objetos comunicacionais para estabelecer as informações e o conhecimento que integram diferentes áreas e particularidades das ciências.

A teoria geral terminológica buscou solucionar uma série de questões, bem práticas para o cenário das produções científicas, da produção e do manuseio de produtos resultantes de técnicas e de tecnologias. Escola wustiana, com participantes de Viena, deram encaminhamento aos trabalhos. Houve um pressuposto marcante: solucionar problemas de sinonímia, polissemia e ambiguidades por meio da normalização. A homogeneidade dessa perspectiva abriu espaço para uma prática centrada na institucionalização das linguagens de especialidades, de acordo com a evolução de teorias e o estabelecimento de conceitos e termos, e na defesa de um planejamento metodológico que formaliza o processo de registro e divulgação.

Para Wüster (1998) o ponto de partida são os conceitos para permitir suas determinações claras e os elos entre eles, além disso, prevê que a Terminologia adote uma posição distinta da Linguística em relação às expressões linguísticas, pois o importante a considerar são as determinações dos conceitos e o ponto de vista essencialmente sincrônico⁸, uma vez que o enfoque recai para uma língua internacional e a prioridade na forma gráfica dos termos. Outro ponto teórico, seria a univocidade que cada termo é atribuído a um conceito, evitando a polissemia e outras dispersões.

É nesse contexto que Cabré (2005) aponta o reducionismo e a uniformização como causas gerais decorrente da teoria geral como parâmetros que restringem as concepções que buscam entender os aspectos comunicativos dos termos diante das variações, dos fatores influentes nas estruturas conceituais e os aspectos socioculturais inerentes às linguagens. Ou seja, implicando em desdobramentos como a da proposta de unidade terminológica globalmente difundida pelo conceito e seu significado causando a redução desse significado se consideramos a inserção em diferentes sistemas linguísticos (social, político e econômico), pois excluiria a dinâmica de construção e uso em vários cenários socioculturais.

Contudo, foi através dos esforços para uniformização que os estudos terminológicos clássicos contribuíram para a elaboração dos padrões e normas nacionais e internacionais concentrados em perspectivas de universalização das ciências e viabilizar a aderência das terminologias enquanto o somatório de termos registrados e controlados segundo o contexto e

⁸ O ponto de vista sincrônico segundo Rauch e Barité “concentre-se em eventos linguísticos que são considerados como ocorrendo em um determinado momento da evolução da linguagem.” (2017, p. 214, tradução nossa)

os aspectos ontológicos e epistemológicos de cada especialidade. Um exemplo disso são algumas contribuições absorvidas pelo Comitê Técnico de Terminologia da ISO (International Organization for Standardization) (Krieger, 2000).

Uma das influências ligadas ao estudo dos tesauros está o entendimento do conceito e suas relações para com o termo, ambos apontados por Wüster (1998) na TGT e no que ele chamou de ‘lexicografia terminológica’, cuja noção concentra-se nas atividades de compilação e estruturação de dicionários e glossários terminológicos. Reintegrando os componentes conceituais e nucleares dessa teoria, compreende-se que “todo conceito, com exceção dos conceitos de objetos individuais, corresponde aos elementos comuns que os seres humanos percebem em muitos objetos e que utilizam como meio de classificação mental (para entender) e, portanto, também para comunicar.

Portanto, o conceito é um elemento do pensamento” (Wüster, 1998, p. 39, tradução nossa). Aproximações conceituais estabeleceram uma nova interface para priorizar os estudos acerca dos conceitos e termos nas investigações e na atuação de vertentes teóricas como a da representação indexical ou temática e das linguagens documentárias. A questão das dispersões, das relações entre os termos e o tratamento lexical focados nas especialidades podem ser considerados vitais para as proposições terminológicas.

Diante da amplitude que ganhou essa teoria, dois cenários de trabalho terminológico são desenhados: o trabalho descritivo destinado à compilação e à análise da denominação conceitual; e o trabalho prescritivo focado na orientação conceitual para garantir a univocidade e a monossemia (Cabré, 2005). A TGT teve um papel no alicerce terminológico, do qual se colocou diante de críticas e de oposições, no sentido de formular outras perspectivas teóricas não abordadas anteriormente.

Outros fenômenos surgiram posteriormente as décadas 70, sendo imprescindíveis reavaliar até que ponto a prescrição normativa e as suas consequências estariam dispostas a observar e a tratar os repertórios técnicos e científicos perante transformações históricas. Diante dessas questões, Krieger contribui com a seguinte reflexão acerca dos aspectos gerais dessa vertente:

À problemática da feição da exterioridade significante dos termos técnico-científicos, soma-se a ideia de que estes são meras etiquetas designativas a serviço da consolidação e divulgação das ciências e das técnicas. No bojo desse pensamento, encontra-se a concepção wüsteriana de que a terminologia expressa conceitos e não significados. Ao contrário destes que são lingüísticos e variáveis conforme o contexto discursivo e pragmático, os conceitos científicos são estáveis, paradigmáticos, universais. Tal concepção está fundamentada num modelo positivista de ciência que ainda compreende a língua científica como lugar homogêneo e transparente que tem

por única função expressar as verdades científicas. Nessa ótica, o valor da unidade terminológica define-se apenas pelo lugar que ocupa na estrutura conceptual de uma especialidade. Explica-se também assim a concepção de que o termo técnico-científico é uma unidade de conhecimento, antes de que expressão linguística (2000, p. 218).

Pode-se reconhecer então que a fundamentação pautada no modelo positivista de ciência confere a TGT uma autonomia diante dos estudos gerais linguísticos e aqueles centrados no léxico geral. Além disso, lhe confere uma afinidade estreita com as ciências exatas e seus postulados universais, já que aqueles permanecem por mais tempo exercendo formas de referências concretas e consensuais para as comunidades de especialistas dos quais fazem parte.

Considerando o apanhado geral, outras teorias foram surgindo no final do século XX com outros objetivos. A Teoria Comunicativa da Terminologia é um exemplo de outro caminho traçado nos estudos e na prática terminológica voltados para perceber problemas na mediação terminológica e com o foco em propostas metodológicas direcionadas aos esquemas de comunicação especializadas, dentre as quais considerando o caráter interdisciplinar da Terminologia e os aspectos socioculturais, econômicos e políticos.

Nessa vertente, Bentes Pinto e colaboradores (2020) no artigo sobre “O léxico das comunidades indígenas do Ceará na designação de doenças: reflexões para a construção de vocabulário controlado” trazem outro olhar concernentes às relações sinônimas e quase sinônimas orientação “USE” indicando a obrigatoriedade pelo uso do termo preferido entre os vários sinônimos tal assertiva não foi levada em consideração pois ‘no processo de comunicação e com vistas à preservação da memória lexical, qualquer palavra pode ser usada’. Portanto, a decisão dos autores foi adotar “USADO TAMBÉM PARA” (UTP), como nos exemplos:

Difruço
 UTP- Defruz
 UTP -Defurço
 NE = “Influenza” [gripe] com outras manifestações respiratórias, devida a vírus não identificado (CID-10)
 Ex: Falta de Fôrgo UTP - Perder o Fôrgo
 UTP - Cansaço
 NE = Dispneia, fôlego curto, ortopneia (CID-10). (BENTES PINTO *et al.* 2020, p. 189)

Diante dessas novas percepções e das contribuições da TGT, possibilidades nas etapas, padrões e outros aspectos terminológicos são estabelecidas na construção de tesouros.

3.2.2 Teoria do Conceito segundo Dahlberg

Ingetraut Dahlberg apresentou contribuições significativas nos estudos de classificação e organização do conhecimento com a Teoria do Conceito, ou Teoria Conceitual Analítica, que viria motivar a fundamentação teórica com um olhar clínico para as unidades representacionais presentes em diferentes tipos de estudo na Ciência da Informação. Influenciada pelos princípios analíticos-sintéticos da Teoria facetada de Ranganathan, pela Terminologia e pela Filosofia, Dahlberg (1978a, 1978b) aponta o conceito como unidade para se compreender aquilo que nomeamos em um conjunto de conhecimento.

Essa teoria se afasta da tradição norte-americana quanto ao seu tratamento técnico em função da etiquetagem centradas nos documentos e seus respectivos assuntos predominantes. O horizonte da Teoria Conceitual Analítica vai além e permite formalizar as relações conceituais presentes nos termos. Toma como premissa as análises conceituais diante de esforços terminológicos. Implica que todo o conceito tem um referente e quaisquer declarações verificáveis designam possíveis propriedades e relações das quais tal referente possa denotar (Dahlberg, 1978a).

Os precedentes e o suporte dessa teoria podem ser acompanhados nas bases filosóficas e linguísticas responsáveis por orientar a organização do conhecimento e as terminologias, dentre elas estão: Lógica, Teoria da Ciência, Epistemologia, Ontologia, Fenomenologia, Aletologia e Metafísica (Dahlberg, 1992). A concepção filosófica consolida a proposta da autora, principalmente, diante da aplicação de princípios e métodos pertinentes aos padrões terminológicos, e mais precisamente na formação definitiva dos termos e suas relações intrínsecas com os conceitos. Ou seja, há uma inquietação com a validade do conteúdo dos conceitos e a função deles para as linguagens de especialidades.

A iniciativa de Dahlberg (1978a) ampliou o diálogo dos conceitos, posterior ao cenário das comissões terminológicas, a padronização internacional, em específico com o INFOTERM e outros trabalhos de Eugen Wüster. A autora buscou encaminhar a discussão conceituais para o âmbito das ciências sociais, pois estas apresentavam uma configuração diferenciada se em comparação com as ciências exatas. Essas últimas bem mais beneficiadas com o advento das normas internacionais.

Assim, retomando, não somente, o melhoramento da comunicação científica, mas apoiando diretamente a consolidação de disciplinas e campos científicos voltados ao social e às humanidades. Portanto, fenômenos e objetos científicos requerem uma construção conceitual e deve-se considerar que “a confiança quase exclusiva dos cientistas sociais no uso de termos

derivados de usos da linguagem comum resulta em uma proliferação extrema dos significados em que as palavras mais comumente usadas são empregadas, produzindo assim uma confusão polissemântica que parece desafiar todos os esforços de normalização” (Dahlberg, 1978a, p. 142, tradução nossa).

Para Dahlberg (1992, 1978b) as linguagens e seus elementos tornam possíveis a fixação do conhecimento, os atos de se comunicar e de designar objetos do mundo. É a partir da identificação dos objetos no tempo e no espaço que a pesquisadora defende os primeiros indícios da denominação conceitual, ressaltando o ponto de vista de Immanuel Kant a partir do momento em que cita que o ato perceptivo dos objetos perpassa pelos sentidos, e a diferenciação entre os conceitos gerais e os individuais que se dá com a captação dos elementos assimilados (Dahlberg, 1978b).

Não é descartada a percepção de objetos advindos do pensamento humano frente a dicotomia existente entre conceitos teóricos (contextos teóricos) e conceitos empíricos (de objetos da realidade). Pois há um processo de integração em uma estrutura social, não sendo uma divisão entre dois lados opostos, mas sim uma descoberta de que os conceitos teóricos incluem os empíricos (Dahlberg, 1978a). A construção de um conceito, mesmo perpassando pela percepção subjetiva do sujeito, está integrada ao espaço empírico e social a partir da relação desse sujeito com o entorno. Assim como também, há possibilidades de um objeto abstrato, como figuras e personagens lendários fruto de um pensamento humano, seja reconhecido como integrante de um determinado grupo, ou seja, o conhecimento compartilhado por esse grupo. O que viabiliza a exploração da dimensão ontológica.

A diferenciação dos elementos constituintes e descritivos evidenciam características e atestam o que é um objeto individual, conforme a sua localização no espaço e no tempo, o que por outro lado informa que seu conceito é individual. Já os objetos gerais seriam aqueles cujos conceitos gerais são empregados fora de uma determinação precisa de tempo e espaço. Um exemplo, é o conceito geral de “homem” e atribuição substantiva que não faz ideia se a entidade está atrelada a um dado período ou uma localização espacial. Contudo, o conceito individual de ‘o primeiro homem a pisar na lua’ esclarece o fato da existência de um homem que pisou na lua em 20 de julho de 1969, fornecendo indicativos de lugar e data.

Assim, é necessário compreender que os elementos dos conceitos são aferidos de acordo com o método analítico-sintético e o reconhecimento de que cada enunciado mostra um atributo predicável do objeto, ou seja, uma característica (Dahlberg, 1978b). Para tanto, o cerne central da questão teórica recai na definição de conceito, como sendo “[...] a unidade de conhecimento que surge pela síntese dos predicados necessários relacionados com determinado

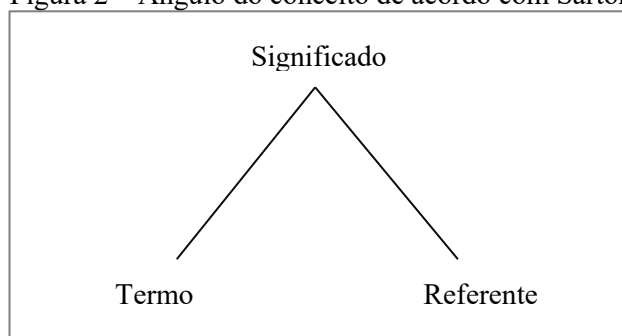
objeto e que, por meio de sinais linguísticos, pode ser comunicado” (Dahlberg, 1978c, p. 12, tradução nossa), em outros termos “[...] um conceito é uma unidade de conhecimento que compreende as características de um referente por um termo ou um nome” (Dahlberg, 1978a, p. 144, tradução nossa). Desse modo, são descritos elementos de composição para tornar possíveis as indagações sobre a natureza conceitual e sua identificação de acordo com os sistemas linguísticos. Há então um processo de compilação dos enunciados, dos quais se configuram como um conjunto de predicados em meio ao elo existente com o símbolo linguístico.

Podemos agora definir a formação dos conceitos como a reunião e compilação de enunciados verdadeiros a respeito de determinado objeto. Para fixar o resultado dessa compilação precisamos de um instrumento. Este é constituído pela palavra ou por qualquer signo que possa traduzir e fixar essa compilação. É possível definir, então, o conceito como a compilação de enunciados verdadeiros sobre determinado objeto, fixada por um símbolo linguístico (Dahlberg, 1978b, p. 102).

Esses enunciados verdadeiros denotam possibilidades analíticas, diferentes daquelas descritas nas normas terminológicas, uma vez que recorrem a leitura e a inferência ontológica e epistemológica para tratar conceitos como aqueles referentes às disciplinas científicas - esses que requerem verificação e validação. Para tanto, Dahlberg (1992) recorre aos filósofos Immanuel Kant e Gottlieb Frege que elucidaram a predicação nas acepções conceituais e afirma que um conceito é elaborado pela predicação sobre o seu referente, sendo uma noção orientada ao referente.

Uma elucidação clarificada dessa teoria, concentra-se no apontamento das seguintes etapas de construção de um conceito: passo referencial, passo predicacional e passo representacional, sendo ambos componentes representados graficamente na proposta de Giovanni Sartori (Dahlberg, 1978a). Na figura 2 conferimos essa ilustração mencionada pela autora:

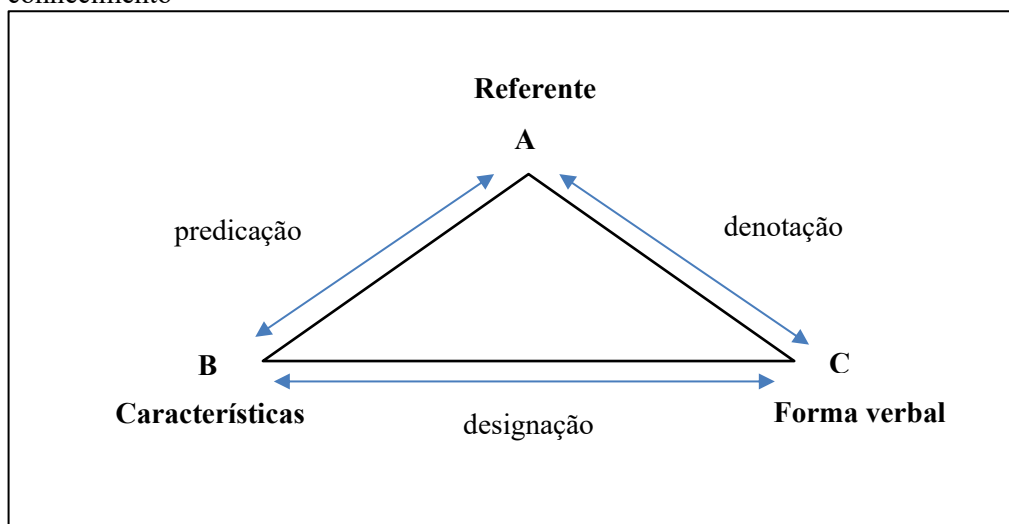
Figura 2 – Ângulo do conceito de acordo com Sartori



Fonte: Dahlberg (1978a, p. 144, tradução nossa)

Com a releitura dos componentes conceituais do ângulo de Sartori constituído pela tríade, significado, termo e referente, Dahlberg (1978a) expõe que esse esquema não considera o “universo de itens”, de acordo com a formação conceitual e o “universo do discurso”, quanto à aplicação, visto que ambos os universos devem ser acionados ao termo e ao referente respectivamente. Aqui a autora menciona um aspecto dimensional importante na condição de comunicação: o universo do discurso. Requer uma atenção também para a socialização do conceito consolidado conforme sua utilização por uma dada comunidade discursiva ou falante. Com essa reinterpretação, Dahlberg apresenta sua proposta no triângulo da Figura 3 abaixo:

Figura 3 – Triângulo do conceito, a representação formal de uma unidade de conhecimento



Fonte: Dahlberg (1978a, p. 144; 1992, p. 66, tradução nossa)

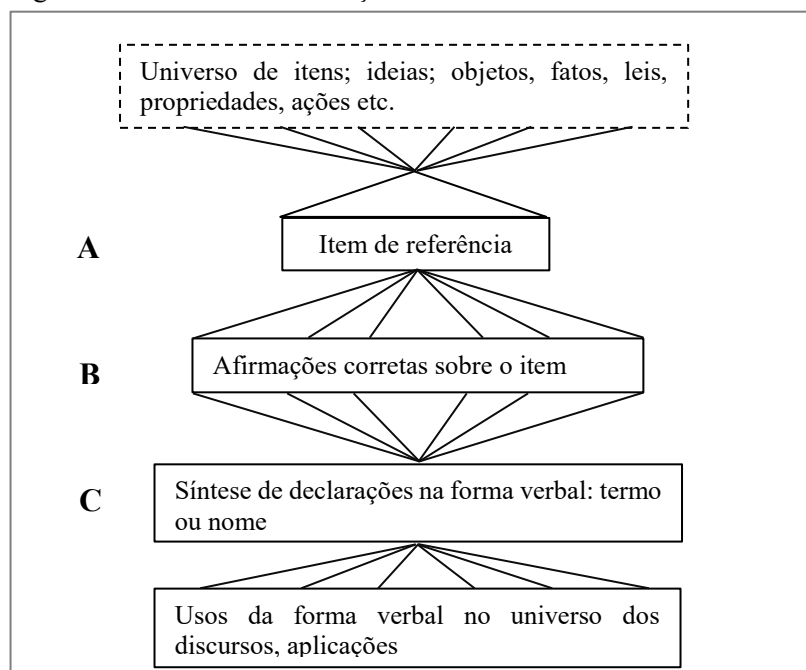
No diagrama acima, um referente (A) possui predicação (enunciados verdadeiros) responsável por apontar as características (B) de tal objeto (referente), dos quais essas características designam uma forma verbal (termo ou nome). Logo, o objeto coração sendo um referente indica um conjunto de predicados - órgão do corpo humano, localizado no sistema cardiovascular, composto por músculos, cavidades e vasos, sua função é bombear o sangue para outras partes do corpo por meio de uma rede de vasos sanguíneos - possuem características (órgão muscular do sistema cardiovascular) e uma forma verbal ‘coração humano’.

A elucidação visualizada na triangulação acima, sinaliza extremidades angulares correspondentes ao referente (o objeto atrelado a sua imagem e sua ideia percebidas pela mente humana); ao significado (tudo aquilo que pode ser qualificado e atribuído como verdadeiro e enunciativo desse objeto); e por último ao termo (unidade que representação a formação de tal

conceito perante o sistema linguístico e comunicado). O foco dado por Dahlberg, concebe uma dinâmica intrínseca entre os componentes de cada etapa de construção de um conceito.

A perspectiva dos elementos descritivos funciona como critérios de avaliação do objeto e as possíveis desambiguações em relação ao sistema linguístico. No exemplo de coração, as características são fundamentais para afirmar que o exemplo do objeto ‘coração’ se refere ao ‘coração humano’ e não a outro objeto que possivelmente possa apresentar a mesma forma verbal. A explicação do triângulo do conceito vale para conceitos estabelecidos e convencionados em um dado domínio ou disciplina. Existem, ainda, aqueles conceitos e termos em formação de designação e definição, dos quais ainda não passaram por eventuais estudos. Entretanto, na Figura 4 abaixo é possível descobrir como ocorre a formação conceitual e a investigação de seus elementos:

Figura 4 – Modelo de construção do conceito



Fonte: Dahlberg (1992, p. 66)

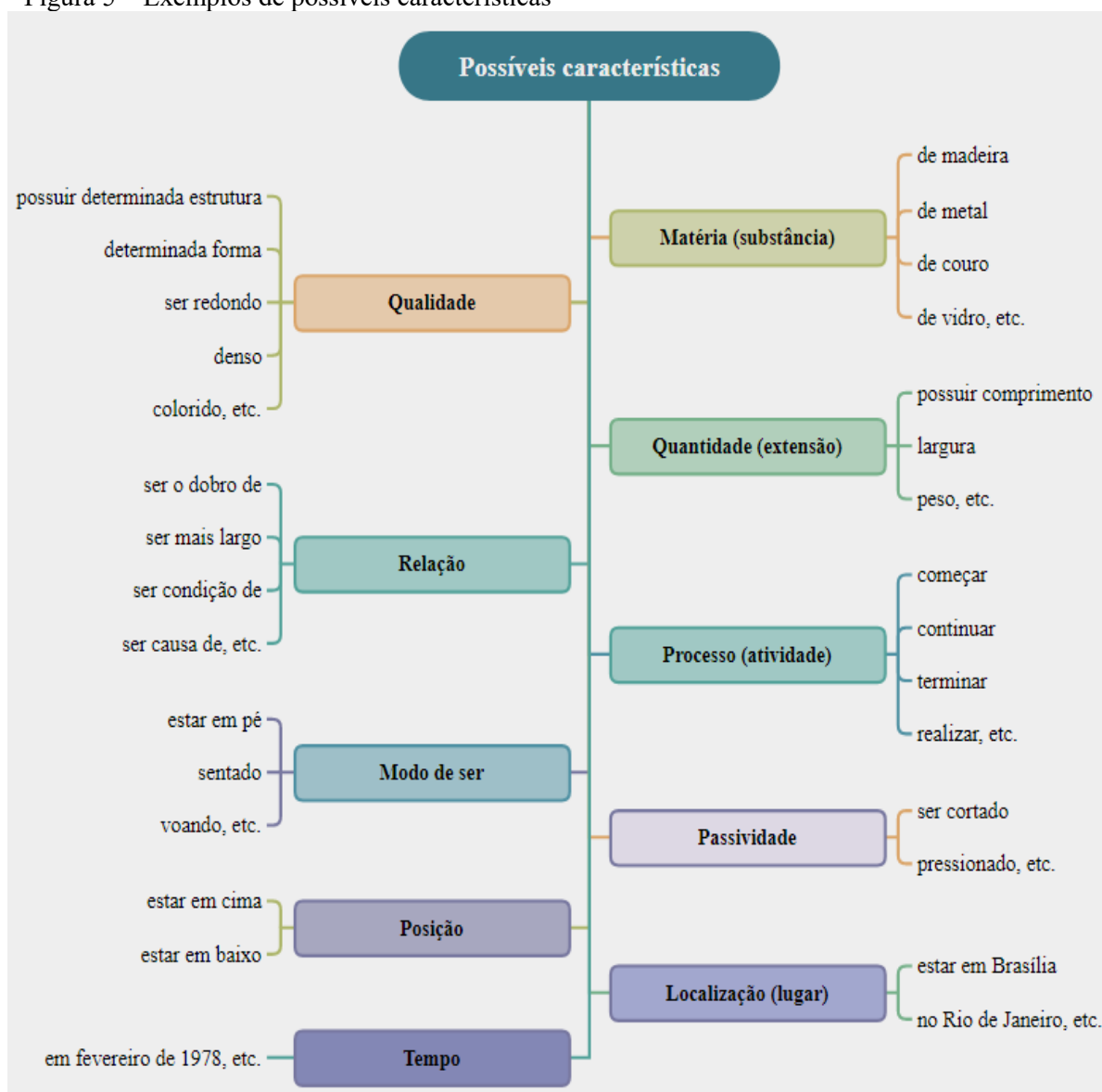
Apoiado nessa noção que percorre o objeto enquanto ponto referencial até a sua forma verbal, encontram-se explicações que satisfazem o uso de um conceito em um determinado contexto e quais aspectos semelhantes e distintos se comparados com outros conceitos. O resgate e a introdução das categorias de Aristóteles servem de parâmetro para elencar as características do objeto de modo a distingui-las, dentre as espécies categóricas estão: tempo; processo; matéria; propriedades; modo de ser (Dahlberg, 1978b, 1992). Mais

precisamente, uma categoria é uma característica mais geral de um conceito e esta pode representá-lo de forma ampla.

Outro ponto fundamental da teoria é a característica, entendida como um componente do conceito e provida a partir de um enunciado acerca de um referente (Dahlberg, 1978a). Características idênticas permitem a relação entre os conceitos, como apontado por Dahlberg (1978b), são exemplos dos conceitos semear, colher, transportar a colheita, preparar o solo, a característica comum: a atividade agrícola. Dividem-se em dois tipos: características complexas, aquelas que possuem mais de uma propriedade; e as características simples, aquelas que possuem uma única propriedade (Dahlberg, 1978b; 1978c).

Na Figura 5 é possível acompanhar exemplos de espécies de características apontadas pela autora e inspiradas nas categorias aristotélicas:

Figura 5 – Exemplos de possíveis características



Fonte: Dahlberg (1978b, p. 103)

Nessa perspectiva, as observações analíticas tendem à descrição das características dos objetos utilizados, tanto para compor os enunciados, quanto para agrupar relações conceituais de acordo com as semelhanças comuns comparadas com outros objetos. Essa colaboração teórica é indispensável para o estudo e estruturação das relações conceituais presentes em classificações e tesouros. Com esse intuito, a Teoria do Conceito buscou interpretar o entendimento do “ser objeto” até a sua concepção, tomada por um conceito e sua relação com os demais, seja de acordo com uma teoria ou um tipo de conhecimento (filosófico, popular, científico).

As relações são consideradas a ligação existente entre dois ou mais conceitos possuidores de características idênticas e podem ser classificados de acordo com essas

características, pertencendo, assim, a um grupo com uma ou mais características comuns. Elas podem ser de dois tipos, um relacionamento quantitativo e um relacionamento qualitativo. A relação estabelecida de modo quantitativo mede a quantitativamente a semelhança entre características que um conceito apresenta, enquanto a do tipo qualitativo são fixadas exclusivamente em subdivisões de acordo com os tipos de conceitos (gerais, individuais) e de acordo com os referentes (Dahlberg, 1978a). Para compreender os relacionamentos, o Quadro 1 sintetiza as colaborações de Dahlberg:

Quadro 1 – Tipos de Relações entre conceitos

Tipo	Descrição	Exemplificação
Relações hierárquicas (gênero-espécie)	Estabelecidas quando dois conceitos diferentes possuem características idênticas e um deles apresenta uma característica adicional, de modo que surge entre eles uma hierarquia.	Árvore Árvore frutífera Macieira
Relações partitivas	Estabelecidas entre um conceito que representa um todo e qualquer uma das partes desse todo.	Navio Proa Casco Convés
Relações funcionais (sintagmáticas)	Estabelecidas exclusivamente na dependência do conceito de processo.	Pinturas – quadros – pintores – compradores de quadros
Relações de oposição	Estabelecidas quando ocorre oposição entre conceitos, sendo possíveis de duas maneiras: contraditoriedade e contrariedade.	Branco - não branco (contradição) Branco – azul (contrariedade)

Fonte: Adaptado de Dahlberg (1978b, 1978c).

Os relacionamentos estabelecidos entre conceitos configuram-se como mecanismo de agrupamento e ligação semântica, uma vez que são pautados na proximidade dos seus significados, origem ontológica ou etimológica de cada um. Pode-se acompanhar pelos tipos que cada ligação é feita mediante a interpretação dos enunciados e dos elementos conceituais descritos (proposições predicativas), além da compreensão de que o conceito indica um determinado termo advindo de um determinado sistema conceitual. O termo aqui trata-se de um componente vital, principalmente, para os estudos dos conceitos no âmbito das linguagens de especialidades e das particularidades do conhecimento humano. Segundo Dahlberg (1978a) o

termo é uma forma verbal que resume o conceito com o propósito de representá-lo nos processos e meios de comunicação.

Reconhecendo as dificuldades de entendimento perante o uso e construção de diferentes origens e funções, a teoria demonstra outros suportes para evidenciar e constatar a ideia e a expressão do conteúdo de um conceito. A definição dos conceitos é compreendida como um desses suportes. Ela é imprescindível para impor os limites da ideia considerando um sistema composto por outros conceitos. Dahlberg (1978b), uma definição é um modo de circunscrever e determinar os limites de um conceito, ou seja, delimitar um conteúdo com uma intenção preestabelecida e ou apresentar um conjunto de características e atributos. Considerando os sistemas de classificações e vocabulários estruturados, a definição implicará na compreensão e cobertura conceitual, além de princípios, dentre os quais, o de especificidade e o de uso do termo.

A amplitude de medição de conteúdo ou da definição de um conceito pode imprimir dois níveis: a intenção e a extensão. Ambos permitem auxiliar no estabelecimento de relações conforme o número ou o detalhamento de características e atributos na definição. A intenção de conceito é a totalidade de suas características, correspondendo a totalidade de conceitos de suas hierarquias mais suas características específicas, já a extensão é a soma total de conceitos especiais e individuais, podendo abranger os objetos reais ou itens, possivelmente, indicados como itens de referência dos conceitos (Dahlberg, 1978a).

3.2.3 Teoria da Classificação Facetada

Frente às classificações bibliográficas, no final do século XIX a Teoria da Classificação Facetada foi aos poucos recebendo notoriedade conforme o reconhecimento das noções analíticas e sintéticas. Sendo relevantes para a composição de esquemas classificatórios firmados na ordenação de livros e outros documentos nas estantes. A Classificação Decimal Universal foi a primeira manifestação, embora os trabalhos de William Charles Berwick Sayers (1881-1960) tenham de fato expressado os aspectos fundantes, baseados em uma metodologia alicerçada na lógica e em outros princípios filosóficos, dos quais se destinavam ao campo da organização do conhecimento, mais precisamente às classificações em bibliotecas (Hjørland, 2013a).

Apesar de William Charles Berwick Sayers ter esboçado a teoria no âmbito das bibliotecas, foi com o seu aluno, o matemático e bibliotecário indiano, Shiyali Ramamrita Ranganathan que a classificação facetada alcançou resultados mais nítidos e adquiriu a

qualidade de teoria moderna da classificação. O próprio Ranganathan comenta do estímulo que recebeu das obras de Sayers na introdução de *Prolegomena to library classification*. Embora Hjørland (2013a) indica a primeira manifestação na Classificação Decimal Universal conforme alguns aspectos da classificação facetada, Campos (2001) afirma que a origem está na *Colon Classification* desenvolvida por Ranganathan para a Biblioteca da Universidade de Madras.

O teórico matemático identificou e procurou canalizar esforços em princípios básicos que garantem a resolução de problemas frente à obsolescência de tabelas classificatórias. Assim considerando os assuntos enquanto unidades a serem investigadas racionalmente diante de um universo de ideias que confrontam o pensamento humano. A preocupação estava em como perceber diferentes entendimentos acerca de um livro, um objeto ou qualquer outra ideia atrelada a um determinado assunto, desse modo corresponderia a estruturação guiada pelo método facetado. Essas noções podem ser identificadas em publicações rangathanianas, destaque para *Colon Classification* de 1933 e *Prolegomena to Library Classification* publicada a primeira vez em 1937.

Com a introdução e o desenvolvimento de uma base sólida e consistente, a Ranganathan sustenta conceitos referenciais para ilustrar que as classificações se iniciam com o entendimento dos assuntos na dimensão do conhecimento, antes mesmo do exercício do método e aplicação de princípios para a estruturação de tabelas. Há espaço para a distinção entre informação, conhecimento e conceito. Abaixo o Quadro 2 aponta aspectos conceituais relevantes para as reflexões aqui observadas.

Quadro 2 – Conceitos introdutórios apresentados por Ranganathan

Assunto	Um corpo de ideias organizado ou sistematizado, cuja extensão e intenção provavelmente se enquadram coerentemente no campo de interesse e confortavelmente dentro da competência intelectual e do campo de especialização inevitável de uma pessoa normal. (Ranganathan, 1967, p. 82, tradução nossa)
Ideia	O produto do pensar, refletir, imaginar etc., obtido pelo intelecto, integrando com o auxílio da lógica, uma seleção aperceptiva e/ou o que é diretamente apreendido pela intuição e depositado na memória. (Ranganathan, 1967, p. 81, tradução nossa)
Conhecimento	A totalidade das ideias conservadas pelos humanos. Nesse sentido, Conhecimento = Universo de Ideias. (Ranganathan, 1967, p. 81, tradução nossa)
Informação	Ideia comunicada por outros ou obtida por estudo e investigação pessoal ⁹ (Ranganathan, 1967, p. 81, tradução nossa).

⁹ Ranganathan (1967, p. 81) menciona a proximidade de ambos os conceitos (conhecimento e informação) relativa à possibilidade de uma sinônima: “*Knowledge and information are sometimes treated as synonyms*”, ou seja, são tratados às vezes como sinônimos.

Conceito (Concept)	A formação, depositada na memória, fruto da associação de percepções — puras e compostas — já depositadas em memória (Ranganathan, 1967, p. 81, tradução nossa).
Universo	É qualquer agregado de entidades sob consideração. A divisão de um universo é o processo de classificar as entidades desse universo em subagregados com referência a uma característica escolhida, colocando entidades semelhantes no mesmo subagregado e entidades diferentes em diferentes subagregados. A divisão também será usada para denotar o resultado da divisão, 7.e., um grupo de subagregados formado por uma divisão das entidades de um universo ou aquele estado de tais entidades em que elas foram divididas. A característica escolhida pode ser chamada de característica de divisão. Exemplo: o conjunto de meninos em uma sala de aula é um universo (Ranganathan, 1937, p. 10, tradução nossa).
Esquema de Classificação	Um esquema de classificação pode ser definido como uma declaração que mostra um arranjo coerente ou filiativo das classes que surgem no decorrer do fornecimento completo de um dado universo. Um esquema de classificação, portanto, [...] envolve três conceitos: (1) Características; (2) Matrizes de classes; e (3) Cadeias de classes. (Ranganathan, 1937, p. 26, tradução nossa.)

Fonte: Ranganathan (1937; 1967)

É importante a atenção dada para a definição de conhecimento, pois o autor eleva o estudo para reconhecer que a “totalidade das ideias conservadas pelos humanos” recorda o fato de que as ideias se originam na mente, ou seja, no pensamento humano, antes mesmo dos objetos que correspondem a fixação desse conhecimento. Os objetos, no caso, podem ser entendidos como aqueles que portam e imprimem tal conhecimento, dentre os quais livros e outros tipos de documentos. Os outros elementos como: o assunto, o conceito e o universo são essenciais para essa teoria.

Conforme a análise de *Prolegomena* e a questão do desenvolvimento do universo de assuntos de acordo com o conteúdo dos documentos, Campos (2001) ressalta Ranganathan como precursor da ideia de que os livros deveriam ser abordados por diferentes aspectos, apontando um grau de complexidade em virtude das distintas formas de encarar um problema e deixando claro que o conhecimento de diversas áreas acontece em virtude do progresso da Ciência. Com o destaque dessa obra, o grupo Classification Research Group (CRG) retomou e deu continuidade ao trabalho original de Ranganathan¹⁰, concentrando pesquisas e estudos de classificação e indexação na vertente britânica e europeia.

O diferencial teórico está no quão dinâmico os assuntos se dispõem a representar e a categorizar o conhecimento. Isso talvez tenha sido um forte aliado para o tratamento de vocabulários mais especializados (terminologias) em detrimento daqueles extraídos da linguagem natural. Broughton (2006b) explica que durante o estabelecimento dessa teoria, o

¹⁰ Ranganathan iniciou seus estudos e atuação na Universidade de Madras na Índia, depois prosseguiu para a University College em Londres, aproximando-se dos pesquisadores britânicos e do Classification Research Group.

grupo do CRG se preocupava com a classificação destinada à organização da mídia impressa (catálogos e fichas), perceberam então algumas vantagens na classificação facetada relativas aos questionamentos na produção de uma ordem linear para assuntos não lineares, ou seja, assuntos compostos. Assim, estavam progredindo quanto ao tratamento de vocabulários compostos e selecionados para esquemas bem mais robustos em termos semânticos.

Outro pilar estrutural é formado com base nos princípios da Teoria Dinâmica do Conhecimento¹¹. Essa perspectiva considera que o conhecimento é parte de um movimento espiral, desenvolvido diante dos processos de compreensão dos fenômenos do mundo e, conseqüentemente, a Espiral do Conhecimento mencionada por Ranganathan implicou na percepção do Universo de Conhecimento e do Universo de assuntos. Com a releitura da Teoria Dinâmica na Teoria da Classificação Facetada, Ranganathan apresenta uma visão filosófica do desenvolvimento do conhecimento, permitindo representar o conhecimento registrado de um determinado momento histórico perante uma interação com a realidade e os diferentes domínios do conhecimento (Campos, 2001; Campos; Gomes, 2003).

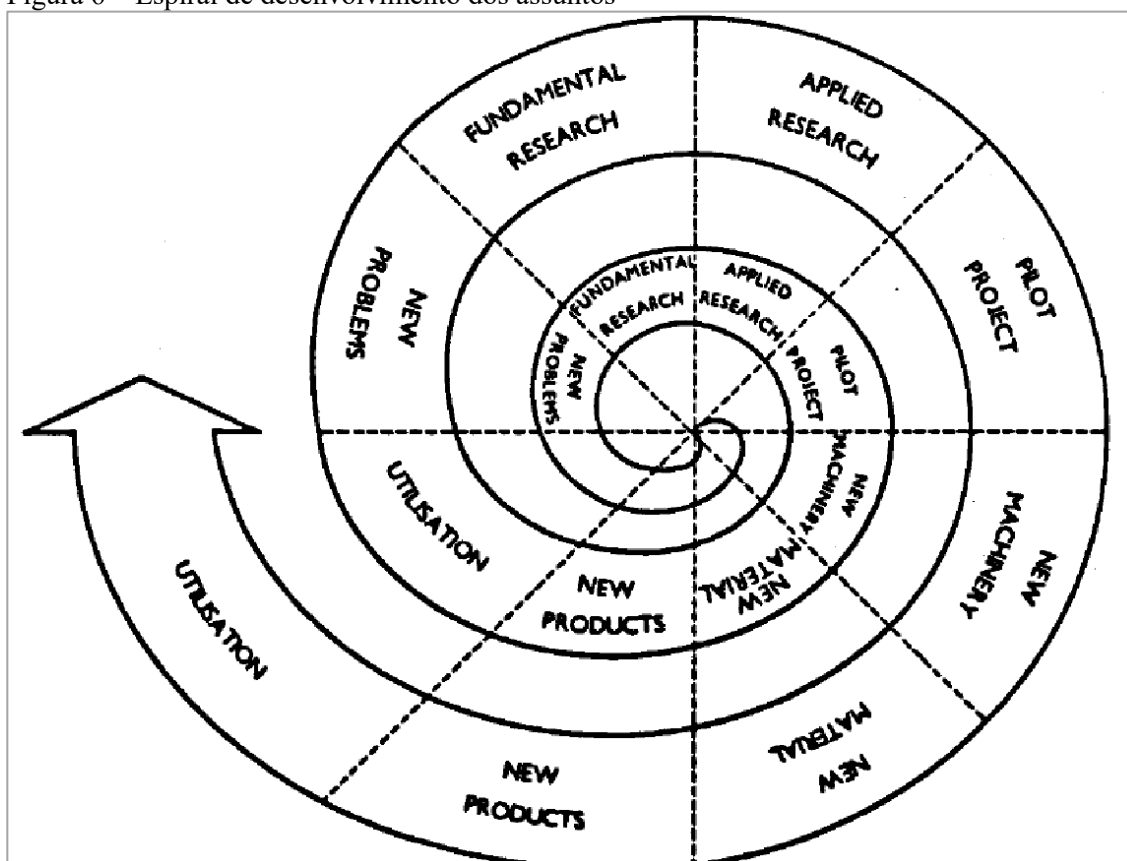
Dessa forma, viabilizou a constituição de esquemas flexíveis que acompanham o ciclo do conhecimento registrado e os demais aspectos determinantes na identificação e estruturação de assuntos. Além disso, a Teoria Dinâmica do Conhecimento serviu para basilar aspectos teóricos para uma Teoria Dinâmica da Classificação Bibliográfica, apresentada nas obras ranganathanianas em postulados, princípios e cânones, cuja terceira edição da *Prolegomena* adota os seguintes planos de trabalho de classificação: Plano ideacional, Plano verbal e Plano notacional (Campos, 2001).

Explicando os planos de trabalho, Ranganathan (1967) não descarta discussões tratadas constantemente pela Filosofia, revelando que a mente humana é o lugar das ideias (plano ideacional) e é através da comunicação firmada na linguagem falada (plano verbal) que as ideias existem na dimensão da qual podemos captar a linguagem escrita (plano notacional). No plano notacional podemos identificar elementos transformadores entre a linguagem falada e escrita, pois imprimem formas de difícil compreensão em contextos distintos, como o caso dos sinônimos e dos homônimos.

A Figura 6 identifica elementos da espiral de assuntos, ilustrada por Ranganathan para explicar etapas pré-definidas que justificam o estabelecimento de um dado assunto:

¹¹ Campos (2001) esclarece que a Teoria Dinâmica do Conhecimento esteve detalhada na segunda edição de *Prolegomena to Library Classification* em 1957 e que este trabalho apresentou uma descrição da Teoria Dinâmica da Classificação Bibliográfica desenvolvida até meados de 1955.

Figura 6 – Espiral de desenvolvimento dos assuntos



Fonte: Ranganathan (1967, p. 374)

A espiral inicia-se com novos problemas, seguido de pesquisa de fundamentos, pesquisa aplicada, piloto do projeto, novas máquinas, novo material, novos produtos, utilização e novamente novos problemas, pesquisa de fundamentos e assim sucessivamente. O ciclo percorre etapas das quais aludem a investigação e o método científico, implicando que um assunto tanto se molda, quanto pode ser investigado de acordo com os processos e agentes dessa trajetória espiral. Ranganathan (1967) percorre um exercício de teorização, expondo questões detalhadas da formação, estrutura e desenvolvimento de assuntos no capítulo P de *Prolegomena*, para então colocar em evidência a proposta de classificação.

Pode-se verificar que os assuntos estão estabelecidos conforme uma série de características que se alteram. Como exemplos, temos os postulados científicos que durante uma determinada época foram considerados inequívocos, porém com os avanços das ciências foram reavaliados e subsídios por outros. O mesmo acontece com um conjunto de conhecimento técnico de artesãos ou costumes de uma determinada época e cultura, que pela não perpetuação ou salvaguarda de registros, são esquecidos e dão lugar a outros.

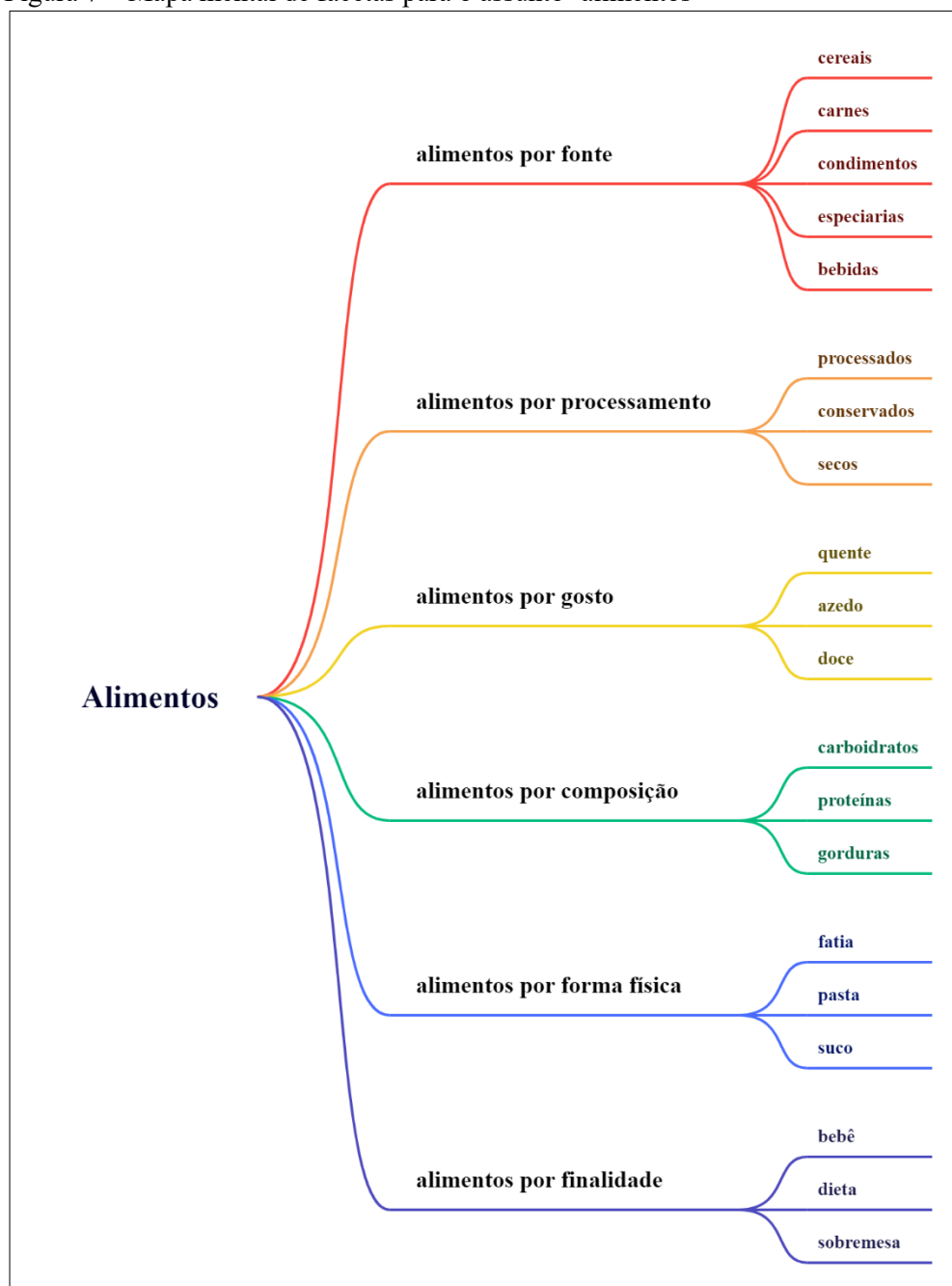
A Teoria da Classificação Facetada concentra-se na análise de facetas, formalizada por meio de conceitos e perspectiva metodológica capaz interpretar e subsidiar esquemas

classificatórios, seguindo também o parâmetro de organização mediante aos sistemas conceituais de áreas e subdivisões do conhecimento. Para Broughton (2002) a análise de facetas é, na maioria dos estudos, entendida de modo restrito, porém pode ser esclarecida considerando-a um processo rigoroso de análise terminológica pelo qual um vocabulário de um dado assunto é organizado em facetas e matrizes, implicando em uma complexa estrutura de conhecimento formada por relações semânticas e sintáticas.

Essa análise facetada representou mudanças em como um determinado assunto poderia ser entendido, considerando algumas características, como o tempo, o espaço ou o tipo de material e processo que um objeto ou assunto poderia ser revelado. Para os classificacionistas anteriores a Ranganathan, os esquemas eram ordenados de acordo com os assuntos representativos da literatura de uma área e em um dado momento histórico, em outros termos, não ocorria a ligação entre o conhecimento e as classificações, mas sim uma relação entre os assuntos dos documentos e as classificações (Campos, 2001).

Nessa análise e exercício classificatório, as facetas revelam que um assunto é notado em diferentes disciplinas ou perspectivas alocadas em um mesmo domínio ou área. O que para alguns sistemas classificatórios isso é inviável. Na figura 7, podemos compreender que o assunto “alimentos” no exemplo complexo apresentado por Vickery (2008):

Figura 7 – Mapa mental de facetas para o assunto ‘alimentos’



Fonte: Adaptado de Vickery (2008).

Cada ramificação apontada corresponde a uma forma de perceber o assunto ‘alimentos’, agrupando conforme características diferentes. Esse agrupamento elencam uma série de conceitos ligados semanticamente a cada faceta desse assunto. A partir dessa noção, as várias facetas relacionadas são diferentes atributos pelos quais um assunto pode ser classificado e um determinado objeto pode ser representado por combinações de seus atributos (Vickery, 2008).

Ranganathan, utilizando inspirações matemáticas, reflete algumas analogias e se refere aos conjuntos infinitos de assuntos indicados através de “facetas”, apresentando-se em dois tipos: as matérias básicas - matérias que não possuem ideias isoladas como componente - e as qualificações de assuntos básicos - as facetas isoladas (Hjørland, 2013a). Foi a partir dessas facetas isoladas que surgiram cinco categorias fundamentais conhecidas pela fórmula PMEST: *Personality, Matter, Energy, Space e Time* (Personalidade, Matéria, Energia, Espaço e Tempo). Assim, o Postulado das Categorias (PMEST ou categorias fundamentais) é constituído como o princípio normativo aceito para organizar um universo de assuntos, sendo o mapeamento de assuntos o primeiro passo para elaboração de um esquema de classificação facetado ou enumerativo (Campos, 2001). O Quadro 3 apresenta cada uma dessas unidades:

Quadro 3 – Categorias fundamentais da análise facetada

Categoria	Descrição	Faceta
PERSONALIDADE	Característica distintiva de um assunto	Mitologia Grega
MATÉRIA	Material físico de que um assunto pode ser composto	Tábua de mármore
ENERGIA	Qualquer ação que ocorre em relação ao assunto	Grande guerra
ESPAÇO	Componente geográfico da localização de um assunto	Atenas (cidade)
TEMPO	Componente geográfico da localização de um assunto	Século 3000 a.c.

Fonte: Adaptado de Hjørland (2013a).

As categorias podem ser visualizadas como tipos específicos de facetas, sendo elas as principais utilizadas para denotar um assunto, ou seja, um conjunto de cinco expressões fundamentais. Nessa perspectiva, Broughton (2002, 2006b) explica que essas cinco categorias e os princípios de categorização implícitos na teoria analítica de facetas podem auxiliar na organização de conjuntos de objetos, termos e conceitos de uma disciplina ou domínio.

Dessa forma, um assunto poderia ser concebido através de cinco componentes primordiais para atestar a existência e o significado daquele assunto perante outros assuntos averiguados durante os procedimentos classificatórios e representacionais. Outra particularidade das categorias fundamentais está na disposição utilizada para estabelecer a ordem das facetas, elencando um ranking de importância referentes às cinco dimensões de cada

assunto de acordo com a concretude decrescente (Hjørland, 2013a). Esse entendimento esteve presente no desenvolvimento de classificações, como a Colon Classification e o Thesaurofacet.

Para entender as categorias fundamentais, assim como alguns aspectos analíticos-sintéticos, Ranganathan apresenta os seguintes conceitos descritos no Quadro 4:

Quadro 4 – Conceitos centrais da perspectiva analítica-sintética facetada

Faceta	Um termo genérico usado para denotar algum componente – pode ser um assunto básico ou um isolado- de um assunto composto, tendo, ainda, a função de formar renques, termos e números classificados. (Ranganathan, 1967, p. 88, tradução nossa)
Característica	É qualquer atributo ou qualquer complexo de atributos com referência ao qual a semelhança ou diferença de entidades pode ser determinada e pelo menos duas delas são diferentes. Exemplos: 1. A altura é uma característica dos meninos, mas a posse do rosto não, embora ambos sejam atributos. A posse do rosto é um atributo compartilhado igualmente por todos os meninos e, portanto, com referência a ele, todos os meninos são como. Assim, não é uma característica dos meninos considerados como entidades. (Ranganathan, 1937, p. 9, tradução nossa)
Renques (Arrays)	Classes derivadas de um universo com base em uma única característica em qualquer etapa do progresso em direção ao seu arranjo completo e organizadas na sequência preferida. (Ranganathan, 1967, p. 61, tradução nossa)
Cadeia (Chain)	Uma sequência de classes composta de qualquer classe e seu universo de 1, 2, 3, etc., até o ponto desejado. (Ranganathan, 1967, p. 63, tradução nossa)

Fonte: Ranganathan (1937; 1967).

Assim conforme os apontamentos de Ranganathan, o assunto é trabalhado conforme a identificação das categorias, posteriormente, a sua composição em facetas, para, então, ser alocados em renques e cadeias responsáveis pela denominação da estrutura central de esquemas como classificações ou um tesouro.

Os princípios rangathanianos estão diluídos detalhadamente na última edição da Prolegomena e foram continuados em pesquisas da Classification Research Group (CRG). Os princípios podem ser identificados nos capítulos, incluídos em algumas partes da obra conforme os planos: o plano de ideias com 14 cânones, 13 postulados e 22 princípios; o plano verbal com 4 cânones; e o plano notacional com 19 cânones (Ranganathan, 1967). Os aspectos metodológicos descritos nos planos evidenciam o modo complexo em que a teoria da classificação facetada trata os assuntos, empregando recomendações precisas para perceber os conceitos, identificá-los no seu universo conceitual e contexto de uso, associá-los conforme suas expressões linguísticas e arquitetá-los em esquemas facetados.

Reconhecemos na teoria facetada a formulação de uma estrutura baseada em renques e facetas que estava influenciando os padrões e recomendações de esquemas e vocabulários estruturados. Desse modo, Campos (2001) destaca os principais cânones dessa teoria utilizados na elaboração e manutenção dos tesauros, dentre os quais estão os cânones direcionados para a formação dos renques: Exaustividade, Exclusividade, Sequência útil e da Sequência consistente; e os cânones para a formação de cadeias: o Cânone da Extensão Decrescente e o Cânone da Modulação.

3.3 Evolução dos padrões aplicados à construção dos tesauros: do *Thesaurus of Engineering and Scientific Terms* a ISO 25964-2

A construção dos tesauros nos primeiros anos de sua implementação esteve particularizada tanto nas edições dos pioneiros na recuperação, quanto nos projetos desenvolvidos de forma restrita ao ambiente das instituições e seus respectivos grupos de pesquisadores. Por outro lado, com as edições, o compartilhamento de artigos e relatórios acerca dos bons resultados somados aos promissores projetos de automatização e sistemas viabilizados por computadores, deram mais ânimo para aprofundar as experiências tesaurais. Nos relatos dos trabalhos apresentados durante a década de 1960 e 1970, já era possível indicar pesquisas para avaliar os tesauros em toda sua completude e financiamentos de vários projetos de caráter avaliativo desenvolvidos pela National Science Foundation e outras agências (Owens; Cochrane, 2004).

Nesse período, os altos custos dos projetos de criação, aplicabilidade e manutenção dos tesauros já eram encarados como um dos principais pontos negativos de sua aderência, esses estão presentes até os dias atuais. Os registros de elaboração de cada etapa e componentes, seguidos de planejamento de infraestrutura e avaliação, eram perspectivas técnicas necessárias aliadas à expressividade dos fundamentos, aspectos conceituais e teóricos. Porém, havia uma descentralização, algo natural, já que os tesauros eram desenvolvidos em contextos particulares com alto financiamento e aos poucos seus desenvolvedores somariam experiências que culminaram nos comitês das diretrizes nacionais e internacionais.

O *Thesaurus of Engineering and Scientific Terms (TEST)*, publicado em 1967, foi o pioneiro na declaração de indicações normativas ao incluir na publicação detalhes das regras que serviram para a sua criação (Clarke; Gilchrist; Will, 2004; Owens; Cochrane, 2004; Broughton; 2006a; Clarke; Zeng, 2012; Clarke, 2016). O *TEST* foi desenvolvido de acordo com as atividades do Projeto *LEX* realizadas pelo grupo *Engineers Joint Council* da agência

governamental Defense Technical Information Center dos Estados Unidos. A edição da publicação de 1967, e na edição de 1969, apresentou o Anexo I - *Thesaurus Rules and Conventions*, reunindo regras e convenções distribuídas em seções. No texto, por exemplo, definem a escolha do descritor - termo utilizado como o preferido - que aparece como os primeiros na lista alfabética em um dos capítulos. O descritor foi atribuído a partir da utilidade dos termos e considerando os seguintes tópicos: a frequência relativa de ocorrência na literatura; a frequência relativa de uso dentro de um sistema operacional; os relacionamentos com outros descritores e precisão; e aceitabilidade científica ou técnica (Engineers Joint Council, 1969). A Figura 8 ilustra um exemplo de termos e sua estrutura no *TEST*.

Figura 8 – Exemplos do *Thesaurus of Engineering and Scientific Terms (TEST)*

SOAP (programming language) 0902	Sociological methodology USE Sociology
UF Symbolic optimum essembly program	Sociology 0511
BT Assembly lenugees	UF Applied sociology
Machine oriented languages	Sociological methodology
Programing lenugees	NT Educetional sociology
Soaps 0703 1111	Industriel sociology
UF †Aluminum soeps	Military sociology
BT Organic selts	Rurel sociology
NT Metellic soeps	Urban sociology
RT—Cleaning egents	RT Authoritarianism
Detergents	Criminology
Dishwashing compounds	Demogrephy
	Dependence

Fonte: Engineers Joint Council (1969, p. 361).

A versão ainda incluiu siglas para cada tipo de termo estruturado, posteriormente, presentes nos padrões internacionais, são eles: USE (*Use*); UF (*Use for*); BT (*Broader term*); NT (*Narrower term*); RT (*Related term*). Em detalhe na Figura 8, os descritores em destaque “*SOAP (programming language)*”, “*Soaps*” e “*Sociology*” apresentam um número para indicar o código de classificação referente ao arranjo sistemático e notacional. As orientações das seções ainda cobrem a indicação de termos expressos no plural e aqueles que deveriam ser descritos somente no singular; o uso de sinônimos ou quase sinônimos; a utilização da sobreposição hierárquica; as especificações para regras sobre referência cruzada entre os termos; além de outros exemplos.

Mesmo apresentando esses aspectos constituintes, o *TEST* inclui brevemente os principais pontos de construção, não detalhando com maior profundidade, por exemplo, o caso das relações dos termos conforme as implicações estabelecidas entre termo e seu conceito. Ou seja, pouco se mencionavam as unidades conceituais como elementares e protagonistas de todo

o processo de estruturação do tesauro. A ressalva está na criação das siglas (USE, UP, NT, RT e BT) pelo *TEST* (Clarke; Zeng, 2012), convencionando um modo mais visível de acompanhar e identificar as relações.

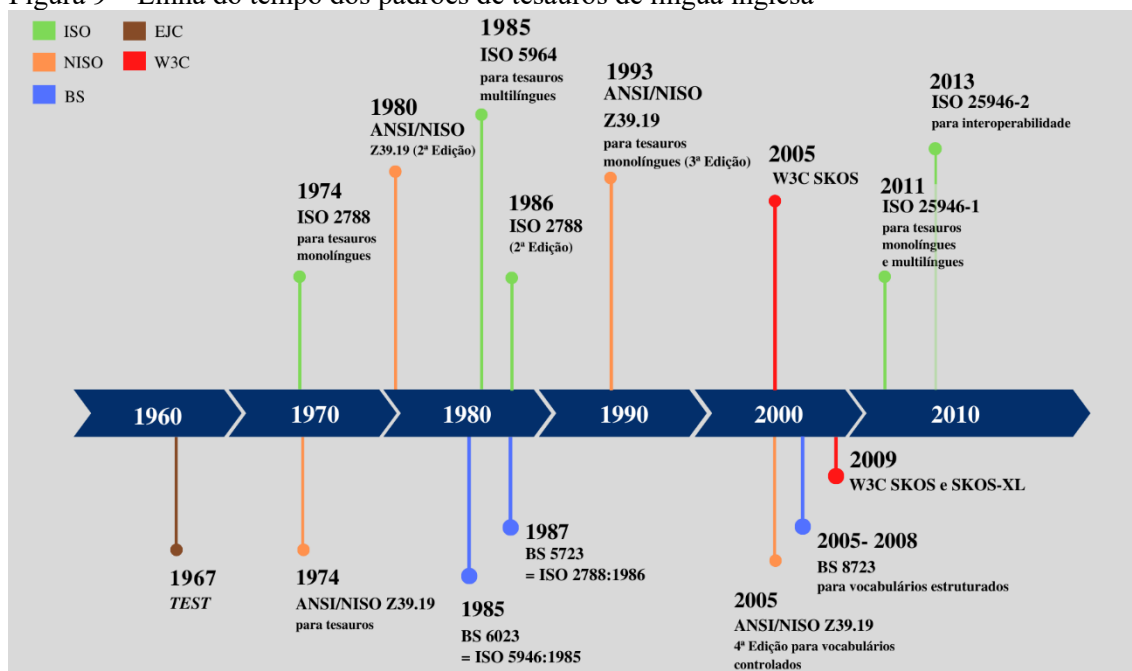
Os princípios e as orientações práticas de construção de tesauros começaram com o padrão alemão DIN 1463 em 1972, seguidas pelo padrão internacional ISO 2788 e pela ANSI Z39.19 desenvolvida nos Estados Unidos pela NISO¹², ambas publicadas em 1974 para tesauros monolíngues, e a ISO 5964 de 1985 para tesauros multilíngues, da qual foi traduzida e formulada para a primeira edição da *British Standards Institution* BS 5723 de 1979 (Clarke; Zeng, 2012; Clarke, 2016). Esses constituíram com a cobertura abrangente, técnica e agregadora de fundamentos teóricos e conceituais que marcaram a história dos tesauros de recuperação, abrangendo o tratamento de vocabulários extraídos de documentos com o propósito de auxiliar na coordenação indexical e nas buscas orientadas para usuários.

Clarke, Gilchrist e Will (2004) explicam que na época de estabelecimento dos primeiros padrões, os tesauros estavam direcionados à aplicação da indexação pós-coordenada e a recuperação de coleções de documentos e dados bibliográficos, contudo, havia a concorrência diante da disponibilidade dos sistemas de computadores para processar as pesquisas de textos complexos e o custo alto para manter um tesauro.

Os padrões tinham o propósito de favorecer os projetos de criação e manutenção, garantindo uma economia em alguns processos, em virtude do emprego de princípios norteadores para ajudar indexadores e demais profissionais. Outro ponto favorável esteve ligado ao compartilhamento e convenção de regras adotadas em nível nacional e internacional, representando a confluência de pontos técnicos basilares. Além disso, fica claro que as orientações deveriam seguir os avanços tecnológicos, principalmente, diante da questão de adaptação aos sistemas e aos requisitos básicos que executam funções essenciais nas redes intranet e internet. Diante desse panorama, Clarke, Zeng (2012) contribuem apontando a trajetória de evolução dos padrões de língua inglesa como ilustrados na Figura 9:

¹² NISO (National Information Standards Organization) é uma organização estadunidense credenciada pela American National Standards Institute (ANSI). A NISO é responsável por promover, editar e publicar padrões técnicos aplicados à gestão de informações.

Figura 9 – Linha do tempo dos padrões de tesouros de língua inglesa



Fonte: Adaptado e traduzido de Clarke e Zeng (2012, p. 22).

Detalhando a linha do tempo, temos a ISO 5964 (*Guidelines for the establishment and development of multilingual thesauri*) ou Diretrizes para o Estabelecimento e Desenvolvimento de Tesouros Multilíngue, publicada em 1985 e, posteriormente, foi adotada por vários países, incluindo Espanha, França e Alemanha, como seus próprios padrões nacionais. Assim, também, no Reino Unido com as nacionais BS 5723 (*Guidelines for the establishment and development of monolingual thesauri*) e BS 6723 (*Guide to Establishment and Development of Multilingual Thesauri*) semelhantes as ISO 2788 (*Guidelines for the establishment and development of monolingual thesauri*) e ISO 5964, respectivamente (Aitchison; Clarke, 2004).

Podemos acompanhar que o padrão NISO dos Estados Unidos e o BS do Reino Unido acompanharam as atualizações propostas pelo padrão ISO internacional. O padrão internacional iniciou recomendações focadas em tesouros monolíngues e edições específicas para tesouros multilíngues, dos quais as edições mais recentes substituíram as anteriores. Os recentes padrões 25946-1 (*Thesauri and interoperability with other vocabularies - Part 1: Thesauri for information retrieval*) e 25946-2 (*Thesauri and interoperability with other vocabularies — Part 2: Interoperability with other vocabularies*) já integram os tesouros monolíngues e os tesouros multilíngues em uma mesma publicação. O destaque está para os recentes padrões do consórcio W3C (World Wide Web Consortium) para o âmbito dos dados e metadados mediados nos sistemas de recuperação e páginas da internet.

Sánchez-Cuadrado, Colmenero-Ruiz e Moreiro (2012) acrescentam a esse cenário as normas espanholas UNE 50-160-90 publicada em 1990 durante a vigência da ISO 2788, a UNE 20-125- 97 de 1997 e as Diretrizes para Tesouros Multilíngues publicada pela IFLA (International Federation of Library Associations and Institutions) em 2009 com o objetivo de complementar a ISO 5964 de 1985 e seu reforço esteve centrado no tratamento da construção de tesouros de forma não simétrica e a ligação entre diferentes vocabulários controlados.

Analisando diferentes requisitos, aspectos e orientações de construção, Somers (1981) fez um estudo crítico das recomendações da International Organization for Standardization ISO 2788 de 1974, a norma British Standards Institution (BS 5723) para tesouros multilíngues de 1979, as diretrizes da UNISIST da UNESCO desenvolvidas no final dos anos 80 e a publicação de Aitchison e Gilchrist, intitulada de *Thesaurus construction: a practical manual* e publicada em 1978. Essa última configura-se como um manual didático que aborda e procura ilustrar exemplos de construção de tesouro com base nos principais padrões vigentes, foi traduzida para diversos países, incluindo uma edição brasileira de 1979.

Compreende-se, a partir da exposição, que os padrões e diretrizes reúnem um material metodológico mais detalhado, especialmente, no tocante de temas e problemas encontrados na preparação de tesouros multilíngues e os fundamentos linguísticos e semânticos associados à estrutura dos termos. O Quadro 5 abaixo mapeia aspectos de construção importantes com base nos tópicos das análises feitas por Somers:

Quadro 5 – Comparativo dos padrões por Somers (1981)

Tópicos	ISO 2788 (1974)	BS 5723 (1979)	Thesaurus Construction (1979)	UNISIST (1980)
Aspectos definitórios de tesouros	- Indica a definição de tesouros tanto em termos de sua função, quanto refletido no aspecto estrutural.	- Reflete o aspecto estrutural ao tratar o tesouro enquanto meio para exibir termos em linguagem de indexação controlada.	- Enfatiza que o valor de um tesouro pode ser estipulado pelo seu desempenho como ferramenta de indexação e recuperação de documentos.	- Indica algumas definições de escopo, semelhantes àquelas do padrão ISO, acrescentando definições acerca do que é um termo, rótulo de nó (indicador de faceta) e idioma de origem.
Aspectos linguísticos	- Indica que é desejável que o descritor adquira o menor número possível de palavras, de preferência, seja apenas uma única palavra; - Apresenta recomendações sobre ortografia e o uso de siglas; - Não faz recomendações sobre a questão de empréstimos e traduções, somente aponta que dever ser utilizadas referências cruzadas entre os dois termos; - Recomenda diretrizes de transliteração.	- Recomenda regras de ortografia; - Apresenta padrões aceitos para diretrizes de transliteração; - Apresenta regras para o uso de siglas.	- Recomenda regras de ortografia; - Apresenta regras para o uso de siglas.	- Apresenta recomendações mais claras sobre empréstimos e traduções, apontando a necessidade do uso de notas de escopo e definições; - Trata do problema especial quando surge em contextos de utilização de vários idiomas acontece o corte de diferentes convenções de transliteração; - Tem atenção especial aos problemas multilíngues relativos aos nomes próprios.
Requisitos formais	- Poucos esclarecimentos acerca do uso de termos compostos e não fornece esclarecimentos acerca das noções de pré e pós-coordenados; - Indica uso de qualificadores para desambiguação de homógrafos; - Aponta a noção “nota de escopo”; - Recomenda que as definições e traduções podem ser utilizadas para desambiguar homógrafos; - Aponta problemas relativos a conjuntos de caracteres; - Recomenda-se o uso de sinais de pontuação.	- Recomendações sobre o uso e a fatoração de termos compostos; - Não fornece explicação acerca dos termos pré e pós-coordenados; - Recomendações acerca do uso da nota de escopo; - Indica o uso dos plurais substantivos; - Informa regras que regem a forma dos termos.	- Detalha e recomendam do uso e da fatoração de termos compostos, incluindo explicativas acerca das noções pré e pós-coordenadas de termos; - Não detalha o termo “nota de escopo”; - Dispõem de regras que regem a forma dos termos, com tabela de regras.	- Menciona o problema dos homógrafos no contexto de interlínguas; - Aponta regras que regem a forma dos termos remetendo o leitor a ISO, embora indique que diferentes convenções se aplicam a diferentes idiomas.

Tópicos	ISO 2788 (1974)	BS 5723 (1979)	Thesaurus construction (1979)	UNIIST (1980)
Relações semânticas	- Apresenta as relações semânticas como tendo uma função indispensável e apresentam as inter-relações entre os conceitos.	- Apresenta as relações semânticas, dentre as quais (subdivididas em sinônimos e quase sinônimos).	- Apresenta as relações semânticas, dentre as quais (subdivididas em sinônimos e quase sinônimos).	- Aponta e discute problemas específicos de tesouros multilíngues de equivalências inexatas de termos, conceitos e relações.
Tratamento multilíngue	—	—	—	- Detalha o estabelecimento de termos equivalentes em diferentes idiomas, além de outras situações de equivalência; - Observa que é viável ser encontradas equivalências exatas, que podem ou não ser mutuamente cognatas, ou expressar o mesmo conceito de diferentes pontos de vista; - Os termos que diferem apenas na conotação devem ser tratados para fins de indexação.
Apresentação e disposição	- Trata de aspectos importantes (conteúdo linguístico e semântico); - Considera o layout físico como importante e dedicam recomendações suficientes para esclarecer sua apresentação; - Indica a elaboração de uma introdução e incorpora recomendações; - Recomenda incluir a exibição alfabética e sistemática; - Descreve, de forma mais ou menos exemplificativa do que recomendatória, as possibilidades de incluir como uma das partes auxiliares do tesouro um display gráfico.	- Trata de aspectos importantes (conteúdo linguístico e semântico); - Considera o layout físico como importante e dedicam recomendações suficientes para esclarecer sua apresentação; - Indica a elaboração de uma introdução e incorpora recomendações; - Indica que o tesouro deve ser organizado principalmente em arranjo sistemático, embora possa apresentar um índice alfabético de apoio; - Recomenda sobre como as informações para cada termo devem ser apresentadas, sugerindo uma ordem específica na qual as diferentes relações devem ser estabelecidas; - Descreve possibilidades para o display gráfico do tesouro.	- Trata de aspectos importantes (conteúdo linguístico e semântico); - Considera o layout físico como importante e dedicam recomendações suficientes para esclarecer sua apresentação; - Recomenda sobre como as informações para cada termo devem ser apresentadas, sugerindo uma ordem específica na qual as diferentes relações devem ser estabelecidas.	- Trata de aspectos importantes (conteúdo linguístico e semântico); - Considera o layout físico como importante e dedicam recomendações suficientes para esclarecer sua apresentação; - Indica a elaboração de uma introdução e esclarecimento acerca do tratamento multilíngue; - Considera diferentes bases para a abordagem alfabética, indicando a problemática multilíngue de variações dispostas em ordem alfabética; - Observa a adoção de símbolos; - Descreve, de forma mais ou menos exemplificativa do que recomendatória, as possibilidades de incluir como uma das partes auxiliares do tesouro um display gráfico.

Fonte: Adaptado de Somers (1981)

As indicações moderam exemplos captados diante da análise de vários problemas apontados por estudos anteriores. A UNISIST, por exemplo, inclui as adequações para o tratamento de idiomas, traduções não cobertas pela ISO 2788 e inova contribuindo com observações acerca da conotação dos termos e dos homógrafos no contexto de interlínguas. A BS 5723, segundo Somers (1981), é supostamente baseada na ISO 2788, sendo que esse padrão britânico foi ampliado para incluir procedimentos com o foco em termos compostos.

Percebe-se que alguns padrões se diferenciam por abordar certos critérios de acordo com o detalhamento, às vezes de modo breve ou, em outros, de modo mais explicativo. Em comum, ambos apostam no tratamento de aspectos importantes referentes ao conteúdo explicativo linguístico e semântico, destacando a identificação de tipos de relações semânticas confrontadas no mapeamento de termos sinônimos e quase sinônimos. Mesmo entendendo que o BS aderiu as ressalvas da ISO, Somers (1981) elencou alguns tópicos que são explorados de forma particular em cada uma dela, por se tratar de adaptações: uma de âmbito internacional e a outra no âmbito da língua inglesa, como no caso do BS.

Outras colaborações podem ser observadas acerca dessas primeiras edições, Clarke e Zeng (2012) esclarece, por exemplo, que embora a ISO 2788 tivesse um objetivo claro de organizar conceitos e suas inter-relações, somente com a edição atualizada em 1974 que se passou a recomendar relações representadas pela sigla de *Broader Term* (BT), ou seja, o termo geral de acordo com a representação em relação a um determinado conceito sendo ele superordenado; e o *Narrower term* (NT), ou termo específico.

Em sequência, a atualização da ISO de tesouros multilíngues 5964 de 1985 declara diretrizes que definem o uso dos conceitos do idioma de origem para traduções direcionadas aos idiomas de destino por meio de regras bem mais explícitas quanto a existência de equivalências exatas, inexatas, parciais, compostas ou não equivalentes (Sánchez-Cuadrado; Colmenero-Ruiz; Moreira, 2012). Essas medidas advêm de situações das quais os tesouros, nesse período, eram adaptados a consulta em diferentes línguas, em uma tentativa de compartilhamento terminológico entre bases de dados que operam para uma gama maior de usuários especialistas.

Contudo, a tradução e o compartilhamento de termos estruturados em tesouros são tarefas difíceis para o trabalho de construção, principalmente, considerando que os métodos de seleção e processamento automático de dados pelos computadores muitas vezes inviabilizam a eficácia do uso dos termos durante os procedimentos de recuperação.

O padrão de tesouros monolíngues ANSI/NISO Z39.19, de 1993 compatível com a ISO 2788 (1986), cobriu com detalhes precisos de certas áreas de controle de vocabulário e se

sobressaiu como o primeiro padrão a lidar com telas e sistemas de controle de gestão (Aitchison; Clarke, 2004). O processamento de vocabulários incluindo a variedade de termos simples, termos compostos e equivalência das traduções, por exemplo, esteve diante das questões semânticas que precisaram acomodar tanto as observações dos profissionais, quanto corresponder aos mecanismos e comandos executados por máquina.

3.3.1 Construção de Tesouros e as perspectivas de interoperabilidade e da Web Semântica

A continuidade dos padrões seguiu um caminho demarcado por mudanças tecnológicas e pela reafirmação dos tesouros enquanto estruturas, não somente vista como instrumento terminológico e especializado, mas sim enquanto uma nova promessa para a organização do conhecimento no mundo digital. Os tesouros impressos não eram os únicos, pois os tesouros a serem arquitetados deveriam auxiliar na representação, classificação e recuperação de diferentes tipos de documentos digitais, dentre os quais imagens, páginas, artigos, mídias audiovisuais. Além disso, os desafios na recuperação não eram mais os mesmos encontrados nas décadas de 1960 a 1970.

Analisando esse cenário que se desenhava a partir da recuperação na web, Garcia-Marco (2014) observou que com o poder e o baixo custo da recuperação vetorial e probabilística, dentre as quais apoiadas por índices de citações como os utilizados pelo Google, ocasionaram uma baixa no uso dos tesouros nas décadas de 1980 e 1990. Nesse período, as recomendações não eram adaptadas para concorrer com motores de buscas e índices criados para a troca e o tratamento de dados na internet. Precisavam então, da ampliação de comandos e das perspectivas de construção dos tesouros que dialogassem com o design e a sistemática da gestão de vocabulários para além dos sistemas de recuperação de informações em intranet.

Aitchison e Clarke (2004) indicam duas tendências principais no desenvolvimento dos tesouros, considerando seus princípios e suas diretrizes, a primeira é a busca pelo acomodação para corresponder a um vocabulário controlado mais rápido, fácil e intuitivo; e a segunda seria o impulso dado pela interoperabilidade que requer projetos de vocabulários para integração em aplicativos *downstream* - como os sistemas de gerenciamento de conteúdo; interfaces de indexação; meta-marcação; portais e mecanismos de pesquisas.

A necessidade de adaptação ao novo panorama, foi evidente no século XXI, além do sucesso da Web Semântica, de Tim Berners-Lee, através a associação de paradigmas da inteligência artificial e da biblioteca com o ideal de criar um cenário modelo que aliasse uma série de inovações, dentre elas, a melhoria na precisão da recuperação na Web e na relevância

da comunicação entre os diferentes sistemas e suas semânticas (Garcia-Marco, 2014). Clarke e Zeng (2012) acrescentam ainda que a era digital animou a criação de muitos vocabulários, dos quais de diferentes tipos, e em algumas vezes trabalhando com tesouros tradicionais (impressos), e, também, trouxe demandas por interoperabilidade para sustentar as atividades dos serviços da web, agregação e troca de dados de tesouros por meio de várias mídias e formatos, além de outras atividades.

Refletindo com base na literatura, a interoperabilidade é a característica que mais solicitou que os padrões e diretrizes de construção dos tesouros fossem de fato reavaliados e reformulados. A interoperabilidade pode ser compreendida como um aspecto primordial de comunicação entre os sistemas para que haja mecanismos integrantes e conexos, permitindo o processamento, o compartilhamento, a leitura e todos os requisitos funcionais mediante as linguagens, padrões, modelos e aplicações. Aitchison e Clarke contribuem no trecho a seguir, informando como os tesouros esteve diante da interoperabilidade e seus desafios:

A interoperabilidade de sistemas é outro aspecto de tornar as coisas mais fáceis para os usuários. Antigamente, depois de consultar o tesouro impresso, começava-se a digitar o (s) termo (s) selecionado (s) no sistema de indexação ou no sistema de pesquisa. Hoje em dia, a ineficiência e os erros potenciais associados à recodificação são inaceitáveis. Os termos devem ser copiados do tesouro eletrônico, copiando e colando ou, preferencialmente, clicando sobre eles. O sistema de pesquisa deve ser capaz de interagir com o banco de dados do tesouro para que os usuários possam navegar durante a pesquisa, talvez selecionando um termo de alto nível e todos os seus termos mais restritos para uma pesquisa exploratória e assim por diante. Essas transações não são difíceis, desde que sejam aplicados padrões para formatos e protocolos de troca de dados. Os requisitos de interoperabilidade assumem outra reviravolta quando o usuário deseja pesquisar vários bancos de dados ou sites, todos usando diferentes tesouros. O mesmo conceito pode ser representado por diferentes termos em diferentes tesouros, e o mesmo termo em diferentes tesouros pode ter significados diferentes. Para permitir a pesquisa entre bancos de dados, alguma forma de mapeamento entre os termos de diferentes vocabulários é necessária. Isso não é tão fácil de conseguir. (Aitchison; Clarke, 2004, p. 16-17, tradução nossa)

Frente a essas questões, podemos compreender que com o advento da interoperabilidade um determinado tesouro pode ser utilizado em uma página da web que já integra outros tesouros, taxonomias e vocabulários, de modo que todos sigam os mesmos protocolos e formatos para promover um melhor desempenho de suas funções de representação, classificação, rotulagem e, conseqüentemente, a de recuperação. Os padrões de tesouros centralizados nas atividades em sistemas de recuperação em ambiente de trabalhos de uma organização ou uma rede interna eram mais adaptados aos padrões anteriores da vigência da interoperabilidade. Porém, quando são incluídos em bancos de dados ou sites da Web, ou precisam atuar em conjunto com outros tipos de vocabulários especializados ou estruturados,

deparam-se com impasses e a ausência real de aspectos construtivos que adaptem os tesauros à Web semântica e ou aos sistemas interoperáveis.

Clarke e Zeng (2012) destacam que muitas das necessidades de interoperabilidade foram incluídas nas recomendações reunidas no Workshop sobre Tesouro Eletrônico, organizado pela NISO em novembro de 1999, influenciando, positivamente, na revisão da ANSI/NISO Z39 e nos novos padrões BS 8723, SKOS (*Simple Knowledge Organization System*) e na segunda parte da ISO 25964. Os critérios focados na manutenção de sistemas de organização do conhecimento por meio de gerenciadores de tesauros automatizados foram incluídos no padrão NISO Z39.19 de 2005 (Sánchez-Cuadrado; Colmenero-Ruiz; Moreiro, 2012).

A amplitude dos novos padrões esteve diante de outras aplicações, como o trabalho feito com os tesauros a partir da necessidade de se relacionarem com ontologias e taxonomias, ou o uso de tesauros para fornecer assuntos em vários esquemas de metadados responsáveis pela descrição e preservação de documentos publicados na web (Aitchison; Clarke, 2004; Broughton, 2006a). Requer nesse intuito, o detalhamento por parte dos padrões ao prezar por recomendações claras que diferenciam as diferenças e semelhanças com relação aos outros vocabulários estruturados e à dinâmica do conhecimento compartilhado.

As tecnologias e ferramentas aplicadas na construção de tesauros implicam em desafios que devem ser ajustados aos aspectos linguísticos, terminológicos, semânticos, entre outros presentes nas recomendações. Por esse motivo, percebe-se uma busca pela harmonia entre métodos, abordagens, tecnologias e padrões. Entretanto, há confusões na utilização e na elaboração dos tesauros com relação aos outros vocabulários e às tecnologias aplicadas.

Nesse cenário, a compreensão das funções e do espoco de aplicabilidade dos tesauros percorre o entendimento de toda infraestrutura tecnológica e a função de outros sistemas de organização do conhecimento, logo resgata os fundamentos basilares que identificam e validam as características desse instrumento. Os métodos utilizados, também, configuram-se como determinantes na aplicação dos padrões, do contrário se não forem aplicados em conjunto com as recomendações, podem implicar no exercício contrário ao trabalho fundamentado e consolidado na evolução dos tesauros.

Os padrões e diretrizes que contemplam, direta ou indiretamente, os tesauros demoram um ciclo para se ajustar às tecnologias e ao trabalho em diferentes ambientes. Suas revisões acontecem em um período, assim como o ciclo de atualização das edições dos sistemas de classificação amplamente utilizados. É pertinente entender que essas orientações por serem apresentadas em inglês e com traduções adaptadas para padrões nacionais em diversos países,

podem, eventualmente, serem vistas como um impasse para a abordagem de problemáticas encontradas em outros idiomas e culturas, já que há um custo e um empenho de reunir profissionais que pensem em casos, soluções e particularidades de acordo com cada país.

Diante dessas observações, os comitês internacionais e nacionais detêm um papel importante e essencial nas orientações mediadas para profissionais atuantes nas atividades indexicais e recuperacionais, quando podemos compreender no trecho de Broughton que atualmente “o tesouro é a ferramenta de indexação preferencial para arquivistas e gestores de documentos, e também para bibliotecários que trabalham com recursos digitais, seja em um serviço de informação local ou em um recurso digital mais amplamente acessível (2006a, p. 6, tradução nossa).

3.3.2 Simple Knowledge Organization System (SKOS)

Nos anos 2000, os padrões foram mais efetivos e se aproximaram, assim como também contribuíram, das estruturas interoperáveis. O consórcio da W3C propôs, nesse meio, as o modelo SKOS (Simple Knowledge Organization System) anunciado 2005 e efetivado em 2009 através de recomendações, fazendo referência aos sistemas de organização do conhecimento (taxonomias, tesouros, esquemas e vocabulários controlados) e suas aplicabilidades perante o conhecimento compartilhado nos ambientes da web.

Segundo Miles e Bechhofer (2009) os sistemas de organização do conhecimento compartilham uma estrutura semelhante e são usados em aplicações semelhantes, a partir dessas aproximações, o SKOS é um modelo que busca tornar explícito as semelhanças e permitir o compartilhamento de dados e de tecnologias em diversos aplicativos. Além disso, possibilitam que os dados sejam interoperáveis durante as aplicações de indexação, classificação e migração almejadas pela Web Semântica. Sánchez-Cuadrado, Colmenero-Ruiz e Moreira (2012) colaboram explicando que SKOS define e implementa um modelo com estruturas básicas de conteúdo a serem compartilhados, permitindo a reutilização e a interoperabilidade.

Podemos reconhecer o papel e os pontos fortes dos tesouros ao visualizar as contribuições do SKOS diante da necessidade generalizada de transformar sistemas e ambientes informacionais da web em espaços de relação e reciprocidade semântica. García-Marco (2016) aponta que diante do espaço de informação global dominado pela internet, os pontos fortes dos tesouros podem ser resumidos como o equilíbrio de seu poder de representação e sua simplicidade, além disso contribui para a unificação de alicerces fundamentais da organização do conhecimento. Para tanto, o SKOS fornece uma ponte para diferentes comunidades da

Biblioteconomia e da Ciências da Informação, das quais desenvolvem projetos e aplicações de sistemas de organização do conhecimento (Miles; Bechhofer, 2009).

O SKOS está intimamente ligado às atividades da Web Semântica e seu comprometimento com os aspectos semânticos no fluxo contínuo das fronteiras entre os sistemas e as páginas da web. O trabalho nesse sentido está sendo desenvolvido pela W3C com o intuito de elaborar e aperfeiçoar protocolos, linguagens, formatos, aplicações e outros padrões, ambos representantes de uma força tarefa em prol de estruturas para auxiliar na catalogação e indexação de informações e dados não estruturados, dos quais impactam diretamente na organização de coleções e modelos de conhecimento.

O modelo aproxima experiências tradicionais que deram certos e as recentes práticas de desenvolvimento e modelagem de dados e metadados. A utilização do SKOS permite que um sistema de organização do conhecimento possa ser organizado e revelado em dados legíveis por máquina (Miles; Bechhofer, 2009).

Ramalho (2015) explica que o SKOS foi uma alternativa de baixo custo que permite a transição de sistemas de organização do conhecimento tradicionais para ambientes na web, funcionando para favorecer uma maior reutilização e interoperabilidade entre vocabulários. Percebe-se que esse modelo não substitui outros padrões e modelos anteriores, mas sim, visa contribuir de modo agregador para adaptar as metodologias e diretrizes consolidadas e difundidas.

Para alcançar seu propósito, o modelo utiliza algumas principais aplicações tecnológicas: as especificações da RDF (*Resource Description Framework*) e a RDFS (*Resource Description Framework Schema*), indicadas pelo W3C para modelagem de metadados; a OWL (*Web Ontology Language*) responsável por auxiliar a RDF nos esquemas da Web Semântica; a linguagem SPARQL (*Query Language for RDF*), também, adaptada à RDF. Para Miles e Bechhofer (2009) a RDFS e a OWL são linguagens que fornecem formas para expressar significados e complementar as informações presentes na Web, funcionando diante de grandes blocos de informações mediante a construção de mapas detalhados de domínios particulares do conhecimento e descrição precisa.

Pastor-Sánchez (2016) aponta que uma das características mais significativas desse modelo está na definição de dois níveis de representação: o nível lexical e o nível conceitual. A descoberta desses níveis de acordo com esse modelo, implica na reorganização das recomendações dos recentes padrões de construção de tesouros. Em outras palavras, sugerem que as orientações sigam de acordo com essas especificações e linguagens do SKOS para promover um ambiente onde a apresentação e gestão de dados e informações acerca de

vocabulários, tesouros e taxonomias, possam ser reutilizadas e introduzidas em projetos de organização do conhecimento. Assim, parte de um vocabulário estruturado de acordo o SKOS possa ser compartilhado por diferentes ambientes, ou uma taxonomia empregada em uma página seja complementada através de termos pertencentes a um determinado tesouros.

Nesse modelo, há rótulos, tags, propriedades, etiquetas e outros elementos empregadas no mapeamento e desenho das estruturas modeladas. As propriedades de etiquetagem, são exemplos desses componentes, proporcionam a atribuição de etiquetas textuais utilizadas na descrição de recursos conceituais (Ramalho, 2015). O processo de descrição que utiliza essa estrutura baseia-se, também, nas relações definidas de acordo com o entendimento dos conceitos. Miles e Bechhofer (2009) complementam identificando que os conceitos presentes no modelo SKOS são baseados em algumas premissas:

- a) Os conceitos podem ser documentados com notas de vários tipos;
- b) Os conceitos podem ser vinculados a outros conceitos do SKOS mediante as propriedades de relação semântica;
- c) Os conceitos são possíveis de serem agrupados em coleções;
- d) Os conceitos podem ser mapeados para outros conceitos do SKOS em diferentes esquemas conceituais.

Os esquemas conceituais são alocados conforme as relações, dentre elas as hierárquicas e as associativas, seguindo os mesmos princípios tradicionais de relacionamentos apontados nos tesouros.

4 ANÁLISE DE DOMÍNIO NA CIÊNCIA DA INFORMAÇÃO

Direcionamentos são dados para caminhos científicos responsáveis por explorar questões gerais da própria Ciência da Informação e corresponder às demandas específicas da atualidade. A noção universal presente na abordagem e nos princípios utilizados na elaboração das classificações bibliográficas, por exemplo, é uma forma de operacionalizar todos os aspectos teóricos e metodológicos dos sistemas de classificações universais e decimais, como a Classificação Decimal de Dewey. Essas abordagens e princípios, por sua vez, são intimamente ligados às diferentes vertentes epistemológicas e do pensamento científico, conferindo aspectos distintos que podem ou não se moldar a cada realidade onde a informação e o conhecimento sejam objetos de estudo e de produção científica.

Por muitos anos, o paradigma físico, discutido por Capurro e Hjørland (2007) esteve presente nas pesquisas e na atuação na CI. Entretanto, outros objetos científicos e de estudo, necessidades e configurações do ambiente e das atividades foram remodelados, solicitando outras perspectivas mais adaptadas à realidade vigente. Nesse sentido, a abordagem destinada às análises de domínios do conhecimento vem se consolidando enquanto proposta para guiar os profissionais da CI em suas atividades, principalmente, a partir da década de 1990, quando os estudos de Organização do Conhecimento (*Knowledge Organization*) têm seus esforços direcionados à interpretação dos diversos domínios e perspectivas epistemológicas (Smiraglia, 2015a).

Há realidades plurais no contexto da OC que se modificam conforme as funções e a predisposição de linguagens, sistemas e tecnologias voltadas às necessidades de uma instituição ou grupo de interesse, das quais requerem estudos e pesquisas cada vez mais adequados ao contexto em que são inseridos.

Nesse cenário, Bibliotecários e cientistas estiveram divididos, levando em consideração os aspectos temáticos do tratamento organizacional: os especialistas que trabalhavam em unidades e bibliotecas especializadas e adquirem competências direcionadas às atividades de interesse e necessidades informacionais seletivas; e os profissionais que são responsáveis pelo tratamento de temáticas gerais e que englobam várias especialidades, como por exemplo, os bibliotecários das bibliotecas públicas (Hjørland, 2017). Portanto, o direcionamento dado para determinada área, especialidade e outras particularidades encaminham as novas discussões e a introdução da Análise de Domínio enquanto proposta paradigmática e norteadora dos percursos metodológicos e epistemológicos.

A gênese do conceito de Análise de Domínio (AD) e a sua noção foram propostos no contexto da Ciência da Computação por James M. Neighbors (1980, p. XV-XVI, tradução nossa), sendo a AD entendida como “[...]a atividade de identificar objetos e operações de uma classe de sistemas semelhantes em um domínio e um problema particular. Uma Análise de Domínio é representada por uma linguagem específica de domínio [...]”. Mais tarde, Prieto-Díaz (1990, p. 47, tradução nossa) propõe um conceito para a Análise de Domínio como sendo o processo no qual “[...]as informações usadas no desenvolvimento sistemas de software são identificadas, capturadas e organizadas com o objetivo de torná-las reutilizáveis”.

No âmbito da Ciência da informação, Albrechtsen (1993, p. 221, *tradução nossa*) apontou que AD envolve investigações científicas ou técnicas, “[...] paradigmas e pontos de vista que representam diferentes interesses de conhecimento em um domínio com o objetivo de construir uma estrutura classificatória para capturar, resguardar e servir a sua transferência e compartilhamento”. Posteriormente, Hjørland e Albrechtsen (1995) trazem reflexões sobre a AD evidenciando que a melhor forma de compreender as informações na CI é através dos estudos focados em comunidades de discurso que são partes da divisão da sociedade do trabalho.

4.1 Pressupostos da Análise de Domínio

Hjørland e Albrechtsen (1995) destacaram-se em apresentar uma proposta mais estruturada do que vem a ser a Análise de Domínio, principalmente, ao argumentarem questões de ordem representacional, organizacional e transdisciplinar na compreensão do conhecimento. Uma dessas questões são os eventuais consensos e dissensos acerca das abordagens cognitivas e sociocognitivas, e dos diferentes caminhos que elas direcionam às práticas e às atuações profissionais. Isso elucida o modo como os profissionais especialistas devem estar capacitados para a dinâmica das especialidades em diversos âmbitos e instituições (Hjørland; Albrechtsen, 1995; Hjørland, 2002a). Pretende, então, se consolidar enquanto alternativa metodológica ao paradigma de processamento da informação na CI (Albrechtsen, 2015).

A Análise de Domínio é um paradigma de multimétodos (Smiraglia, 2015a), responsável por compreender uma base de conhecimento (Hjørland, 2002a; Smiraglia; 2014), constituindo-se em um conjunto de princípios meta epistemológicos e estratégias de caráter social que busca a produção de conhecimento informacional de maneira consistente e holística, retratando os atores sociais e polos epistêmicos (González de Gómez, 2013). Esta pode ser visualizada em duas etapas principais de acordo com Hjørland (2002a): a etapa de aproximação

de um domínio específico e suas premissas ontológicas, epistemológicas e sociológicas; e a segunda com a aplicação de um conjunto metodológico que tem o intuito de analisar o domínio.

Conforme Hjørland (2002a), seu período de estabelecimento, no final dos anos 90, reflete no resgate histórico das tradições epistemológicas e filosóficas influentes na Ciência da Informação e na Biblioteconomia. Ainda em outro momento, Hjørland (2004b; 2005) aponta o Empirismo, o Racionalismo e o Positivismo como sendo as principais tradições. Já seu surgimento esteve atrelado às críticas direcionadas à abordagem cognitivista do processamento da informação:

Desde o início, pretendeu-se apresentar uma alternativa metodológica ao paradigma do processamento da Informação na Ciência da informação, que se preocupava com desenvolvimento de modelos para recuperação de informação (IR) e indexação automática. Apesar, ou talvez, por causa do uso generalizado do termo "Análise de Domínio" na formação e pesquisa em CI, existem diversos entendimentos e relatos de sua história e significado. A Análise de Domínio parece ter adquirido flexibilidade interpretativa (na terminologia da SCOT, *Social Construction of Technology*). (Albrechtsen, 2015, p. 560, tradução nossa)

Compreende-se, nesse relato, uma inclinação para uma nova percepção metodológica alicerçada nas observações informacionais que estão para além da interação humano e máquina. E, principalmente, entende os domínios como uma esfera cuja matéria é a própria natureza e unidade formada pelos componentes do conhecimento, assim como a estrutura de um átomo e suas partículas fundamentais. Essa forma de perceber uma unidade se estabelece como uma orientação ou reorientação para os estudos de fenômenos e vertentes epistemológicas pouco exploradas pela CI e próximos à evidência da noção de construção social do conhecimento.

Na definição de AD, podemos assimilar que ela é tratada como um paradigma da CI:

O paradigma analítico de domínio é, primeiramente, um paradigma social que considera a CI como uma das muitas ciências sociais, fomentando [...] perspectivas psicossociais, sociolinguísticas, sociológicas do conhecimento e sociológicas da ciência no contexto da Ciência da Informação. O paradigma domínio-analítico é, em segundo lugar, uma abordagem funcionalista, com o intuito de entender as funções implícitas e explícitas da informação e da comunicação, e de delinear mecanismos subjacentes ao comportamento informacional a partir desta visão. Em terceiro lugar, é uma abordagem filosófico-realista, tentando constatar as bases da CI por meio de fatores que sejam externos às percepções individualístico-subjetiva dos usuários em oposição, por exemplo, aos paradigmas comportamentais e cognitivos (Hjørland; Albrechtsen, 1995, p. 400, tradução nossa).

A primeira noção expõe e ratifica às origens da CI e o entrelaçamento com outras áreas como a Sociologia e Psicologia, pois estas relações estreitas têm permitido absorver as

influências teóricas e epistêmicas fundamentais como o Pragmatismo e o Empirismo, para então, revisar a trajetória científica e abordagens amplamente discutidas nas últimas décadas. A segunda e a terceira noção delimita o caráter funcionalista e filosófico-realista reintegrados a compreender melhor os aspectos científicos de tal abordagem, o que direciona, assim, o debate para a abordagem sociocognitiva, ao invés de seguir o debate do cognitivismo.

Hjørland (2004b) revela que esse apoio é de caráter empírico e foi encontrado na visão sociocognitiva na Psicologia. É relevante salientar que Hjørland e Albrechtsen, 1995, Hjørland (2002^a) e Smiraglia (2015a) apontam a AD como paradigma e abordagem teórica, contudo, recentemente, Hjørland (2017) utiliza, também, o termo “teoria”. Essa discussão não está clara na literatura, mas acompanhando a parte das principais publicações, observamos se tratar de uma proposta paradigmática.

Por conseguinte, Amorim, Viana e Medeiros (2019) observam que a Análise de Domínio tinha o objetivo, durante o início da sua fundamentação e consolidação, de adentrar nas novas perspectivas hermenêuticas, na teoria crítica, no pragmatismo e no pós-estruturalismo. O que admite não somente auxiliar no aporte metodológico dos estudos e pesquisas, mas também, enveredar no estudo da própria estrutura epistemológica da CI e da Biblioteconomia, elucidando aspectos de uma metateoria discutida por Smiraglia (2015a).

Hjørland, Albrechtsen (1995) citam alguns autores fundamentais e conceitos nos estudos da AD:

- a) Saracevic em 1975 - com a visão do sujeito e a visão de literatura do assunto como essenciais nos estudos epistemológicos;
- b) Taylor em 1991 - com os “Ambientes de Uso da Informação”;
- c) Mann em 1993 - com “o assunto específico ou modelo de disciplina”;
- d) Wilson em 1968 - com a importância dos estudos de comunicação do indivíduo na especialidade e nos grupos pertencentes às disciplinas.

Diante das investigações da informação, do conhecimento e sujeitos diante da realidade social. Gonzáles de Gómez (2013) recorre a Habermas para elucidar o conhecimento do ser social enquanto ponto de partida para compreender a perspectiva de domínio. Ou seja, perceber o meio social e a interação entre os sujeitos é um dos primeiros passos para compreender a proposta da Análise de Domínio. Por outro lado, faz referência à influência social e às origens sociais diante dos processos históricos que culminaram na institucionalização, nos organismos e nas atividades do conhecimento (Berger; Luckmann,

2014). Consequentemente, é no âmbito da dinâmica social que a produção do conhecimento científico e outros tipos de conhecimento se efetivam.

Hjørland e Albrechtsen colaboram no seguinte trecho:

A abordagem de domínio de analítica reconhece que os domínios discursivos compreendem atores, que têm visões de mundo, estruturas de conhecimento individuais, preconceitos, critérios de relevância subjetivos, determinados estilos cognitivos etc. Em outras palavras, há uma interação entre as estruturas de domínio e conhecimento individual, uma interação entre o indivíduo e o nível social. Para a nossa principal tese, o ponto de partida é o conhecimento-domínios, disciplinas, ou comércios, e não os indivíduos e, especialmente, não as maquiagens mais biológicas, fisiológicas e psicológicas dos indivíduos. Os indivíduos devem ser vistos como membros de grupos de trabalho, disciplinas, pensamento ou comunidades de discurso etc. (1995, p. 409, tradução nossa)

O esforço está em refletir acerca dessas questões a partir dos atores, das linguagens, tradições e práticas, das quais são categorias analíticas de pesquisa que a Análise de Domínio busca investigar, seja diante da prática institucional científica, seja a partir da imersão nas dimensões social, política, econômica e cultural. Smiraglia (2015a) recorda que precisamos da compreensão empírica sobre as bases de conhecimento em seus contextos.

Por esse motivo, esse paradigma social elucida questionamentos acerca da natureza e das possibilidades aplicadas à produção e ao compartilhamento de conhecimentos, lançando que “[...] o ponto de vista sujeito-conhecimento implica no ponto de vista orientado ao domínio em contraste com o ponto de vista individual” (Hjørland; Albrechtsen, 1995, p. 401, tradução nossa). Em outros termos, segundo o pensamento centrado no domínio, a subjetividade atrelada a percepção e a assimilação pelo sujeito não são suficientes para denotar o conhecimento, já que esse último é notado como integrante de um ciclo social, onde os sujeitos compartilham suas experiências de mundo, seus pensamentos, seus aprendizados e, então, formalizar acordos e contribuições sejam por meio de modelos, teorias, linguagens, discursos, conceitos e tantas outras manifestações.

A AD institui metodologias, embasadas nas abordagens funcionalistas e filosófico-realista, como guia para pesquisadores em diversas linhas de pesquisas na CI e na Organização do Conhecimento. Isso porque a visão centrada no domínio é sugestiva em função da adoção e da sistematização de metodologias para reconhecer ou conhecer aspectos epistemológicos, características e estruturas de conhecimento, além de percorrer fluxos informacionais e a comunicação mediante aos processos e às relações sociais.

Smiraglia (2015a), Smiraglia e López-Huertas (2015) e Hjørland (2017) apresentam o desenvolvimento de pesquisas e estudos tanto aqueles que explicam a abordagem analítica de domínio, quanto aquelas que aplicam esse paradigma a partir de estudos

bibliométricos, cientométricos, críticos e históricos fomentados na CI, na Biblioteconomia, nos estudos de Organização do Conhecimento e, também, presentes em diversas outras áreas, como a física e a mídia social (Smiraglia; López-Huertas, 2015). Albrechtsen (2015) relata, ainda, a contribuição da Análise de Domínio nos estudos acerca da classificação, além das perspectivas metodológicas esboçadas, anteriormente, em modelos de recuperação de informações e indexação automática.

As pesquisas de Smiraglia (2015a; 2015b) aprofundam quais metodologias são utilizadas em estudos referentes a AD. Podemos com Albrechtsen (2015) notar que o domínio é algo a ser explorado, não algo que esteja pronto. Mesmo considerando que parte da pesquisa analítica retome aquilo que já foi registrado e comunicado nas comunidades e nos meios acadêmicos, o processo de análise acontece ao reunir e permitir descrever um panorama amplo com todos os seus componentes fundamentais, como uma espécie de encaixar peças de um quebra-cabeça.

O caráter multimétodo pode ser visto como possibilidades combinadas, das quais o pesquisador de domínio poderá utilizar ferramentas metodológicas associadas para alcançar melhores resultados, ou focalizar em um tipo específico de pesquisa de acordo com métodos e princípios específicos. A Análise de Domínio é bem esclarecida no sentido prático e teórico a partir das seguintes abordagens aplicadas na CI (Hjørland, 2004a; 2002a):

1. Produção de guias de literatura;
2. Produção de classificações especiais e tesouros;
3. Pesquisa em indexação e recuperação de especialidades;
4. Estudos empíricos de usuários;
5. Estudos bibliométricos;
6. Estudos históricos;
7. Estudos de documentos e gêneros;
8. Estudos epistemológicos e críticos;
9. Estudos terminológicos, LSP (linguagens para fins especiais) e estudos de discursos;
10. Estudos das estruturas e das instituições de comunicação científica;
11. Análise de Domínio nas ciências da cognição e inteligência artificial.

Embora Hjørland tenha enumerado e dividido essas possibilidades, outros autores já contribuíram com possíveis acréscimos de outras abordagens. Dentre essas novas propostas

temos: a semântica de bases de dados, por Smiraglia (2015b); a abordagem centrada no domínio desenvolvido dentro e através do processo de planejamento, projeto e construção de um Sistema de Organização do Conhecimento (SOC) específico por Albrechtsen (2015); e o princípio de proveniência enquanto uma abordagem de Análise de Domínio para a arquivística, proposto por Guimarães e Tognoli (2015).

É importante esclarecer que as 11 abordagens correspondem aos segmentos identificados na literatura da OC e da CI, cujas iniciativas promovem uma releitura e novas descobertas. Os estudos críticos na AD na CI, por exemplo, demonstram a revisão de paradigmas e conceitos imprescindíveis para o atual contexto da comunidade científica, principalmente, por elucidar epistemologias decoloniais, estudos de diversidade de gêneros e multiculturalismo. Já na abordagem “Estudos das estruturas e das instituições de comunicação científica”, propicia investigações cerca da temática da comunicação científica, levando em consideração o domínio como objeto central de estudo, ou seja, também retoma que a comunicação científica e seus modelos fornecem subsídios para entender como um domínio se estabelece, em termos de políticas, meios, canais e sujeitos.

A dificuldade de visualizar e propor instrumentos para perceber os elementos e a amplitude de determinados problemas são pontos a serem levantado aqui, visto que apontar para um estudo de um determinado domínio poderá se configurar como um raciocínio de investigação científica e, conseqüentemente, em quais métodos e quais embasamentos científicos podem ser utilizados. Hjørland e Hartel relata isso na pesquisa qualitativa quando explicam o que seria essa descoberta acerca de um domínio:

[...]no contexto da Análise de Domínio deve começar com um estudo interpretativo de alto nível sobre um assunto ou comunidade de interesse. Uma exigência inicial é descobrir os interesses subjacentes, às diferentes concepções daquela área e, então, negociar uma definição ideal do domínio. Nesse processo, o pesquisador cogita várias noções contemporâneas do domínio, bem como suas histórias recentes, antes de chegar a uma conclusão sobre a essência e seus limites (Hjørland; Hartel, 2003, p. 242, tradução nossa).

A contar essas implicações, a Análise de Domínio desperta uma série de ponderações essenciais durante a execução de uma pesquisa: os limites, os conflitos e as proposições de uma área ou especialidade, a interação permitida entre os níveis formação intelectual e o ciclo de conhecimento – a considerar aqueles implícitos e os explícitos - além das questões referentes à institucionalização cognitiva e institucionalização social nas áreas científicas, como bem apresentadas por Whitley (1974) em *Social processes of scientific development*.

Ainda nesse contexto, Martins, Moraes (2015, p. 80) explicam que “[...] a AD vê o universo do domínio como um ambiente em que a informação é constituída socialmente dentro de um contexto próprio com comunidades que determinam e delimitam seus discursos em torno de temáticas definidas, bem como institucionalizam suas relações disciplinares e interdisciplinares”.

Por um lado, favorece a adoção de um aparato metodológico, por outro, focaliza em três dimensões para emergir a base de um conhecimento. Hjørland e Hartel (2003, p. 239, tradução nossa) elucidam as dimensões em três tipos, conforme aspectos teórico-conceituais: “(1) teorias ontológicas e conceitos sobre os objetos da atividade humana; (2) teorias epistemológicas e conceitos sobre o conhecimento e as formas de obtê-lo [...]; e (3) conceitos sociológicos sobre os grupos de pessoas preocupados com os objetos”. Adiante podemos entender esses três tipos através de dimensões interligadas e componentes de um determinado domínio.

4.1.1 Dimensão Ontológica

A premissa ontológica explora a definição do domínio e os conceitos que determinam e justificam sua origem e sua existência. Esta é alicerçada na Ontologia que por sua vez relaciona-se à Metafísica e às investigações filosóficas (Hjørland; Hartel, 2003), buscando responder aos questionamentos e descrever as partes que compõem um certo objeto. Entende-se Ontologia como “a ciência que é dos tipos, das estruturas de objetos, propriedades, eventos, processos e relações em todas as áreas da realidade” (Smith, 2004 *apud* Hjørland; Hartel, 2003, p. 239, tradução nossa).

Assim, a explicação de termos e seus conceitos dá subsídio a composição do trabalho na AD. Essa dimensão viabiliza o senso crítico e a interpretação, considerando que para determinar um domínio deve-se haver consenso entre os sujeitos e discursos responsáveis pela definição da unidade ontológica e, consequentemente, sua semântica (Smiraglia, 2015a). A unidade ontológica fortalece a garantia de um estudo detalhado e esclarecedor, pois lança releituras e entrelaçamentos das diversas abordagens e teorias elaboradas por uma ciência ou qualquer segmento da sociedade.

Hjørland (2004b) cita as considerações de Thomas Kuhn (1970) acerca das ontologias estarem implicadas em teorias e paradigmas, e não podemos pensá-las livremente como estruturas arbitrárias, pois estas não são construtos científicos confinados indefinidamente. Nesse pensamento, são visíveis que, em exemplo, os parâmetros científicos

se modificam e desenvolvem no decorrer da história sob a forma de teorias, paradigmas e outros fundamentos, reconhecendo que “o mundo fornece resistência a nossas conceituações sob a forma de anomalias, ou seja, situações em que fica claro que algo está errado com as estruturas dadas ao mundo por nossos conceitos. Desta forma, a visão de Kuhn pode ser interpretada como (pragmática) realista” (Hjørland, 2004a, p.4, tradução nossa).

O entendimento do que se tem de um domínio pode ser esboçado pela dimensão ontológica ao esclarecer o que é objeto de pesquisa, qual ramo ou divisão do conhecimento ele se encontra e sua composição. Admite as particularidades e as diferenças de tal objeto em detrimento de outros, ou seja, distingue entre os demais ao apontar definições claras.

A presença de estudos ontológicos é recorrente no estudo das representações e da organização do conhecimento e em tanto outros na CI, assim em específico, quando são retomados pelas ontologias tidas como sistemas de organização do conhecimento. As ontologias, nesse caso, são aplicadas no contexto de organização e representação, principalmente, em comunidades científicas, como a Ciência da Computação, Ciência da Informação e Medicina (Carlan; Bräscher, 2011). Elas agregam relacionamentos entre conceitos e são compostas por termos oriundos de um domínio específico (Hovy; 2002), sendo utilizados sistemas de computadores e no âmbito da Web Semântica.

Aqui na perspectiva do paradigma da AD, Hjørland e Hartel (2003) recorrem ao cerne das questões ontológicas na Filosofia, considerando o levantamento de questões acerca dos tipos e da natureza de objetos de estudo que beneficiam as estruturas da ciência. Logo na CI, pode-se nutrir os diálogos que visam bases e vertentes que tratam de conceitos como informação, documento, regime de informação, *subject* e *Information literacy*, entres outros tratados também no âmbito das ciências sociais, como o caso das representações sociais.

Nesse ponto de vista, González de Gómez (2013) faz menção à questão ontológica do domínio da Saúde, exemplificando com uma breve genealogia a importância de localizar e analisar uma pluralidade de paradigmas científicos, aspectos éticos, aspectos culturais e seus embates para definir o que é esse domínio e formalizá-lo por meio do seu escopo e sua abrangência. Deixando claro, que essa reflexão trata o domínio como unidade ontológica e semântica ao permitir a explicação de termos e suas relações. Dentre os termos, destacam-se: “áreas”, “domínios”, “elementos”, “campos”, “objetos”, “problemas”, “assunto” (“substância”/ “diferenciação substancial”) e “tópicos” (HJØRLAND; HARTEL, 2003).

Na AD, o domínio é considerado por Smiraglia como: “[...] um grupo com uma base ontológica que revela uma teleologia subjacente, um conjunto de hipóteses comuns, consenso epistemológico sobre abordagens metodológicas e semântica social” (2012, p. 114,

tradução nossa). Dias (2015, p. 8) sintetiza comentando que um domínio “[...] pode ser considerado uma área do conhecimento ou um determinado campo de especialidade”. Esse conceito é fundamental para o estabelecimento da premissa ontológica, pois estabelece uma proximidade do pesquisador diante do objeto (domínio) e da linguagem terminológica do qual está inserido.

Assunto, tópicos e elementos são termos que fornecem um quadro comparativo para monitorar o percurso de desenvolvimento do domínio. Dahlberg (2009) revela passos para a identificação de conceitos e termos e sua importância nos estudos da organização do conhecimento. Assim, nas ações analíticas esboçadas aqui, torna-se indispensável uma aproximação com a rede conceitual, visto que para Hjørland (2017) o processo de compreensão de domínios é uma espécie de espiral hermenêutica e o conhecimento do pesquisador imprime uma maneira intelectual de se estudar um domínio. Por consequência, os conceitos indicam a coerência e a concretude do estudo ontológico durante a pesquisa.

4.1.2 Dimensão Epistemológica

A premissa epistemológica investiga as estruturas, as tradições, os paradigmas e as teorias que fazem parte do percurso histórico de disciplinas, linhas de pesquisas, especialidades e grupos de interesse comum. Adentra, então, em questões concernentes aos discursos e seus argumentos. O interesse atrelado a descoberta das estruturas epistêmicas torna claro a proximidade da Análise de Domínio das preocupações surgidas a partir da objetividade científica e da busca por diálogos que ensejem identificar aquilo que se caracteriza e se sobressai em termos de conhecimento de uma dada comunidade.

Observando o intuito dessa premissa, Smiraglia (2011) afirma que a AD instrumentaliza o estudo da evolução do discurso mediante os pólos teóricos de uma comunidade de pesquisa, elevando alguns aspectos, dentre os quais as fronteiras culturais e geopolíticas. Assim, as comunidades devem ser vistas como parte dessa interação teórica que articula a construção de conhecimento através de diversos meios e vias de comunicação, como os periódicos, eventos e quaisquer atividade de reunião contínua de grupos, assim como os processos e resultados da institucionalização científica. Portanto, a leitura analítica incide na trajetória de consolidação epistemológica presente em teorias, paradigmas, modelos e abordagens compartilhadas e desenvolvidas em percursos demarcados por uma série de fatores históricos e sociopolíticos, além daqueles relativos ao próprio desenvolvimento técnico e científico.

Epistemologia é o estudo do conhecimento; ou seja, é o estudo daquilo que é conhecido. Epistemologia também é a divisão da filosofia que investiga a natureza e a origem do conhecimento. Portanto, a epistemologia é a ciência do conhecimento; a investigação autoconsciente do conhecimento. Por meio de uma investigação rigorosa, a epistemologia nos diz de onde vem o conhecimento. Os problemas centrais da epistemologia são a definição de conhecimento e os meios de sua aquisição. (Smiraglia, 2013, p. 2, tradução nossa)

Embora os pressupostos ontológicos aprofundem a existência de um domínio, esses não podem ser de todo modo suficientes, uma vez que requer um plano de fundo constituído por alicerces discursivos com base na pluralidade de teorias do conhecimento. Assim a principal tarefa da epistemologia é “[...] a reconstrução racional do conhecimento científico, conhecer, analisar, todo o processo gnosiológico da ciência do ponto de vista lógico, linguístico, sociológico, interdisciplinar, político, filosófico e histórico” (Tesser, 1995, p. 92). Nessa reflexão, a dimensão se aproxima bem mais da gênese e da natureza esotérica das ciências do que da realidade social, pois reconhece que os sujeitos são os cientistas e pesquisadores dentro dos seus ambientes institucionais e em relações sociais restritas em seus grupos de pesquisa e trabalho.

Reconstruir esse percurso é algo que diante da produção e organização do conhecimento torna-se um modo viável de enxergar o domínio. Hjørland e Hartel, comentam essa questão quando:

A produção e a organização do conhecimento não se referem apenas à adição de novos elementos a uma classificação pré-estabelecida. À medida que o conhecimento se desenvolve e evolui, a visão das estruturas do mundo e as relações entre os diferentes conceitos mudam simbioticamente. Partes do mundo que antes eram consideradas desconectadas podem repentinamente tornar-se fortemente relacionadas. Desse modo, as teorias ontológicas mudam conforme as estruturas conceituais e sociais (por exemplo, pelo desenvolvimento de novos campos interdisciplinares) (Hjørland; Hartel, 2003, p. 244).

Isso permite abranger a existência dos elementos construtivos dessa dimensão. Portanto, são intrínsecas as teorias ontológicas e a premissa epistemológica, pois revelam que além da essência do objeto (domínio) há uma trajetória marcada por características epistêmicas responsáveis por tecer a unidade ontológica.

Hjørland e Hartel (2003) apresentam termos responsáveis por localizar as questões epistemológicas, dentre os quais temos: abordagens- metateorias- paradigmas - regimes (por exemplo, regimes de tratamento) - escolas (de pensamento e pesquisa) - sistemas (de pensamento e pesquisa) - tradições (acadêmicas) - tendências. Logo, a descoberta ou a

redescoberta desses termos eleva a compreensão de como se efetiva a estrutura do conhecimento no segmento científico.

Um dualismo é também esboçado no estudo analítico de domínio (Hjørland; Albrechtsen, 1995; Hjørland, 2002a), ou seja, na existência das principais escolas pensamento epistemológico: o racionalismo e o empirismo, fundamentadas na matemática e na lógica, e nas ciências naturais, respectivamente (Smiraglia, 2015a). Assim, a explicação e a estrutura do conhecimento entre as ciências estiveram historicamente conduzidas por essas escolas, de modo que a “[...] orientação matemática chega facilmente a encarar o pensamento como a única fonte de conhecimento, o filósofo provindo das ciências naturais estará inclinado a considerar a experiência como a fonte e o fundamento de todo o conhecimento humano” (Hessen, 2012, p. 55). Ambos os segmentos epistêmicos pautam várias noções, modelos e paradigmas científicos, e requerem olhar atento quando se analisa domínios e seus contextos.

Dentre as diferentes epistemologias, Hjørland e Hartel (2003) destacam: Construtivismo social, Racionalismo crítico, Teoria crítica, Ecletismo, Empirismo, Anarquismo epistemológico (Paul Feuerabend), Epistemologia feminista e epistemologia de ponto de vista, Hermenêutica, Historicismo, Filosofia marxista da ciência, Teoria do paradigma (Thomas Kuhn), Fenomenologia, Pós-modernismo e pós-estruturalismo, Positivismo, Pragmatismo, Racionalismo e Epistemologia social. Cada segmento tem um papel no desenvolvimento discursivo e corresponde à configuração que determina como se dá a validação e o modo como a comunidade observa e descreve o objeto e seus fenômenos.

Hjørland (2002b; 2005) aponta o Empirismo, o Pragmatismo, o Realismo, o Historicismo como as principais epistemologias influentes na CI. Nesse ponto, pode ser identificada a colocação enquanto ciência social aplicada cujos aspectos teóricos foram direcionados pelas principais vertentes das ciências sociais e os percursos demarcados por fatores históricos e filosóficos das ciências. Por intermédio dessas observações podemos visualizar quais são concepções do fazer ciência - processo e disciplina- e consequentemente, compreender a noção de domínio.

Smiraglia (2015b) resgata alguns apontamentos de Hjørland e Albrechtsen (1995), demonstrando que as mudanças epistemológicas acarretam direcionamentos em determinadas ramos da CI, como no caso da epistemologia pragmática e sua influência em contextos históricos, sociais e culturais da Classificação do Conhecimento. Assim como o afastamento das abordagens racionalistas da análise facetada seguindo para o encontro das dimensões universalmente presentes no conhecimento e direciona-se para abordagens não documentais ou baseadas em intertextualidade - abordagens empíricas - à descoberta de estruturas de

conhecimento o (Smiraglia, 2013). Ou seja, a averiguação de algumas características perpassa pela evolução epistemológica, como podemos perceber por meio da evolução dos paradigmas e abordagens na CI a partir do seu histórico epistemológico e, também, filosófico, já que o tecido epistêmico foi possível a partir das contribuições filosóficas.

4.1.3 Dimensão Sociológica

Demarcar as fronteiras das quais o domínio esteve ou se localiza representa a imersão na natureza social, uma vez que parte dos questionamentos que levaram ao aprofundamento da Análise de Domínio resgata a necessidade de olhar para pessoas e grupos enquanto unidade, ao invés de visualizar indivíduos separadamente. Logo, na premissa de cunho social se trata de uma leitura sociológica para identificação dos atores sociais e suas relações diante da construção ontológica e, também, epistemológica do domínio (González de Gómez, 2013).

Considerando a Análise de Domínio na perspectiva sociocognitiva, podemos compreender que sujeitos participam de grupos e compartilham tanto suas problemáticas e necessidades informacionais, quanto experiências adquiridas. Os grupos têm uma relação com os indivíduos que os integram por intermédio da “[...] internalização de signos e símbolos produzidos culturalmente e a forma como os processos cognitivos são mediados por significados culturais, historicamente e socialmente construídos” (Hjørland, 2004a, p.3, tradução nossa). A concepção de grupo então é formalizada no paradigma analítico de domínio através das comunidades discursivas, um conceito que revitaliza a experiência de notar o caráter social da informação.

A Análise de Domínio, portanto, não concebe os usuários em geral, mas os vê como pertencentes a diferentes culturas, estruturas sociais e domínios do conhecimento. Os produtores, intermediários e usuários de informações estão mais ou menos conectados em comunidades que compartilham idiomas, gêneros e outras práticas de comunicação tipificadas. Existem diferentes distâncias semânticas entre os agentes. (Hjørland, 2004a, p.3, tradução nossa)

A pluralidade de comunidades advém desse cenário, demarcado por Hjørland, como o ambiente onde acontecem o compartilhamento de idiomas e outras práticas, sendo complexo percorrer as atividades dos grupos a partir das suas particularidades enunciadas por meio das relações sociais. Um norte para esse trabalho, será recorrer aos termos: “disciplinas (de pesquisa, de ensino); subdisciplinas; comunidades de discurso, comunidades epistêmicas,

profissões, especialidades, sistema social da ciência e uma variedade de coletivos da vida cotidiana (ou seja, "*hobbys*", "amadores", "entusiastas") ” (Hjørland; Hartel, 2003, p. 242, tradução nossa).

Nessa perspectiva, Bourdieu (2004a), ao tratar dos elementos característicos das ciências sociais, colabora destacando o reconhecimento das características e do panorama do "trabalho de objetivação" científica. Esse trabalho nos permite discernir que "construções sociais correntes" são representações e que em um dado contexto podem ser subjugadas e regularizadas em padrões e protocolos, pois as representações que se fazem no campo da ciência nada mais são do que convenções utilizadas para compor os domínios.

Mas o que faz a especificidade do campo científico é aquilo sobre o que os concorrentes estão de acordo acerca dos princípios de verificação da conformidade ao 'real', acerca dos métodos comuns de validação de teses e de hipóteses, logo sobre o contrato tácito, inseparavelmente político e cognitivo, que funda e rege o trabalho de objetivação. (Bourdieu, 2004a, p. 33)

Retomando algumas reflexões e questionamentos sobre a ciência, as comunidades e os seus discursos no âmbito social, podemos reconhecer as construções socioculturais a exemplos a filosofia, a história, a ciência, a arte, a literatura, assim como a existência de um percurso histórico para cada uma delas enquanto resultante dos processos de sociais de construção do conhecimento científico e o conhecimento não científico.

Recorremos a Thomas Kuhn (2005) que situa o fenômeno das revoluções científicas, ressaltando a concepção de comunidade científica da qual pressupõe a organização da ciência por intermédio de uma atividade social. Essa busca por entender a realidade e viabilizar o desenvolvimento do conhecimento perpassa pela história social, de indivíduos e suas trajetórias dentro e fora da instituição ciência e suas estruturas. Considera-se, por exemplo, que a comunidade científica percebe a objetividade que o domínio compartilha.

Um dos pontos principais da premissa sociológica é o diálogo com a Sociologia do Conhecimento e a Sociologia da Ciência. Compreendo a partir de Hjørland e Albrechtsen (1995) que a Sociologia do Conhecimento fornece estudos basilares para interpretar e observar as comunidades e como o conhecimento é produzido, registrado e compartilhado, por exemplo.

Diante do exposto podemos considerar um domínio não somente do ponto de vista particular de disciplinas e campos científico e de interesse, porém também, como produto das relações sociais mediadas a partir das vias e meios informacionais e da construção social do conhecimento. Isso promove a concepção ligada ao reconhecimento das estruturas básicas que sustentam as diferentes perspectivas e abordagens do conhecimento, logo cooperando em

iniciativas que visam o desenvolvimento de sistemas de organização e representação do conhecimento mais precisos e coerentes.

A partir dos diálogos preliminares dos autores, é possível retornar a gênese e as características dos domínios por meio das premissas ontológicas, epistemológicas e sociológicas. Esse tríptico é associado a uma série de atividades que vão desde estudos de análise e compilação de literatura e documentos até a visualização e a interpretação de todos os elementos coletados.

É importante destacar os esforços da Análise de Domínio em desenvolver alicerces teóricos e conceituais necessários para a CI, visto que a sistematização metodológica dada pelo caráter multimétodos viabiliza articular um quadro de métodos e técnicas em diversas combinações e para inúmeras realidades identificadas durante a pesquisa, principalmente, aquelas de cunho empírico. Outro importante destaque é aquele dado às comunidades e o âmbito social. Assim, é possível esclarecer que a Análise de Domínio pode encaminhar a CI aos horizontes da complexidade em que os indivíduos se deparam, em seus cenários de estudo, trabalho e nas relações de construção e disseminação de conhecimento.

4.2 Domínio, comunidades discursivas e discurso

Diante da necessidade de aprofundar cada dimensão e associá-las à prática metodológica como forma de conhecer as ações e efetivar os processos empíricos, visualizam-se categorias conceituais necessárias para a compreensão da essência e da abordagem AD. Nesse ínterim, o conceito de domínio levanta questionamentos acerca do objeto, da forma e de como essa abordagem pode ser consolidada.

Desde as primeiras impressões do paradigma por Hjørland e Albrechtsen (1995), a ideia de domínio se mostrou complexa e, para algumas interpretações, pouco clara. Hjørland (2017) recorda o questionamento levantado por Mai (2008) de que o conceito de domínio não havia sido definido claramente. Esse é crucial, pois a abordagem é gerada na assimilação desse conceito central. Nesse intuito, Mai (2008) propôs uma definição contextualizada nos estudos de indexação centrados nas abordagens de domínio, nas quais o domínio pode ser uma área de especialização, um corpo de literatura ou grupo de pessoas trabalhando em uma organização.

Outras definições foram levantadas: Smiraglia (2012, p. 114, tradução nossa), clarifica a noção, considerando que “[...] um domínio é um grupo com uma base ontológica que revela uma teleologia subjacente, um conjunto de hipóteses comuns, consenso epistemológico sobre abordagens metodológicas e semântica social”. Nesse sentido, se uma hipótese de

domínio, após um estudo analítico aplicado, não indicar uma base ontológica associada ao consenso epistemológico e semântica social, essa não será considerada como domínio. Embora, essa definição imponha condições fixadas nesses pontos característicos, é viável retomá-las à luz das dimensões da Análise de Domínio e verificar que ambas são a existência de vários elementos descritivos fixados para compreender um fenômeno ou objeto (esse supostamente um domínio). Por outro lado, não restringe a ligação entre uma base do conhecimento estabelecida em sua dinâmica social e epistêmica.

Situando esse conceito nos estudos das linguagens documentárias, Lara (2002, p. 135) complementa: “o domínio é um campo especializado do conhecimento (áreas do saber ou de atividades) expresso por uma língua de especialidade (língua utilizada pelo domínio e caráter rizada pelo uso de meios de expressão linguística particulares)”. A autora, aqui, evidencia o domínio como algo advindo do conhecimento especializado, aquele cuja instituição principal está nas terminologias, por emitir discursos e construtos lançados nos processos institucionalização e nos meios de comunicação articulados pela linguagem de especialidade e padronizada. Essa definição não engloba outros tipos e origens do conhecimento, pois está centrada na segmentação do conhecimento técnico e científico.

Salientamos que a definição do domínio, dentro do estudo da Análise de Domínio, é uma tarefa complexa, pois utiliza uma série de pressupostos e redes de conceitos, dos quais ambos são inseridos em conjunturas epistemológicas e ontológicas de natureza interdisciplinar, solicitando averiguar os limites das ciências, áreas, campos, disciplinas e das atividades do pensamento humanos. Além disso, situam-se em períodos históricos, em expressões linguísticas e socioculturais próprias de qualquer objeto de estudo situado no tempo e no espaço.

Visamos explicar de que modo podemos apreender um domínio demonstrando sua função, forma e o gênero dentro dos estudos de Análise de Domínio, Tennis (2003, 2012) expõe uma proposta para definir um domínio e fornece exemplos. Primeiramente, o autor reconhece que um domínio pode ser uma área de especialização ou um sistema de pessoas e práticas com uma linguagem comum (Tennis, 2003). Essa definição está de acordo com a noção de Mai (2008) e vai ao encontro do pensamento de que um domínio deve ser explicitado e definido antes mesmo do início de uma pesquisa.

Como instrumento para definir um determinado domínio, Tennis (2003) apresenta as áreas de modulação, os eixos, e os graus de especialização como parâmetros viáveis na captura de aspectos ontológicos e epistemológicos, situando por intermédio de escopo, do alcance (extensão, exclusão, rótulo, especialização) e do propósito os limites consideráveis para saber forma objetiva desse objeto de estudo. O autor ainda destaca os tipos de Análise de

Domínio: a instrumental é direcionada ao desenvolvimento de sistemas de organização do conhecimento e na elaboração de um determinado produto; e a descritiva é centrada na evidência de constituição de um domínio a partir dos propósitos e interesses dos pesquisadores (TENNIS, 2012).

Entretanto, a perspectiva de Tennis (2003, 2012) diverge dos argumentos apontados por Hjørland (2002a, 2004, 2008, 2017) no tocante a definição e a forma de estabelecimento conceitual de um domínio. Tennis (2003, 2012) é muito mais objetivo e categórico quando pauta eixos e módulos para prospectar um esboço conceitual, não há indícios mais acurados para aquilo que ele acredita ser o conceito de domínio, apesar de demonstrar como fazer para chegar a uma definição de um domínio e a característica do tipo de pesquisa que prossegue aplicando a Análise de Domínio.

Diante das tensões em 2017, Hjørland publica o artigo "*Domain Analysis*" na edição do volume 44 de número 6 da *Knowledge Organization*, trabalho que também está publicado na Enciclopédia da ISKO. Nele, Hjørland resgata o tema e analisa a literatura das últimas três décadas, esclarecendo muitas divergências acerca dos conceitos e elucidando pontos importantes para a consolidação desse paradigma.

O primeiro ponto conceitual a destacar é que a questão se inclina para as referências dadas ao objeto de estudo da Análise de Domínio e aos distintos tipos de pesquisas e formas de apresentação do contexto do qual um domínio se insere. Para tanto, Hjørland (2017) cita a definição de Shapere, em 1977, de que um domínio é um conjunto de informações e é em si uma hipótese que pode ser ou não rejeitada, solicitando uma resposta para tal indagação. Assim, de acordo com o autor, “em particular, se o problema requer uma ‘teoria’ como resposta, o domínio constitui o corpo total de informações que deve, idealmente, ser explicado por uma teoria que resolva esse problema” (Hjørland, 2017, p. 440, tradução nossa).

O destaque para uma “teoria” elucidada de um lado as reflexões de Thomas Kuhn (2005) acerca das mudanças similares notadas pelos cientistas na dinâmica das ciências, e essas resultam na invenção de novas teorias, considerando que nem todas as teorias são paradigmáticas, mas ambas estão diante de um processo de revolução e avanço científico. Percebe-se que as revoluções científicas estão na adequação às modificações tanto do contexto de cada vertente de pensamento, quanto na aderência de novas experiências, por isso podem ser comparadas as transformações revolucionárias de um determinado cenário social e possam emergir em novas maneiras de conceber certos conceitos e formulações, até mesmo, caracterizar-se como mudança paradigmática. Nessa questão, é possível identificar os domínios

como núcleos e unidades a serem estudadas do ponto de vista analítico conforme o pensamento de Kuhn.

Hjørland (2017, p. 439, tradução nossa) expressa que “o domínio por outro lado, é uma especialização na divisão do trabalho cognitivo que é teoricamente coerente ou socialmente institucionalizado [...], os domínios não são divisões prontas do mundo, mas são dinâmicos, em desenvolvimento e dependentes da teoria”. Ao mencionar a coerência, a institucionalização social e o caráter dinâmico, podemos notar uma proximidade com o conceito defendido por Smiraglia (2015) de que o domínio tem uma base ontológica e uma semântica social concebida a partir do conhecimento expressado na linguagem e nos discursos compartilhados por um grupo de pessoas.

A explicação de Hjørland (2017) acerca dos argumentos de Shapere, revela que os domínios são fenômenos e para o seu entendimento é necessária uma investigação cuidadosa e profunda, além disso podem descobrir diferentes perspectivas teóricas. Esse conceito, entende ainda, que o domínio é algo em construção, dinâmico e que percorre um processo evolutivo. No caso de um domínio novo que ainda está se estabelecendo, as dimensões epistemológicas e ontológicas, somadas à observação das relações e interesses sociais de seu grupo, denotam seus aspectos conceituais responsáveis pela sua definição. Diante dessas argumentações, o autor sintetiza uma definição intermediada e agregadora das interpretações de outros autores.

Um domínio é um corpo de conhecimento, definido social e teoricamente como o conhecimento de um grupo de pessoas que compartilham compromissos ontológicos e epistemológicos. Os domínios são frequentemente disciplinas acadêmicas, mas também podem ser, por exemplo, *hobbys*[...]. Do ponto de vista da Ciência da Informação, Biblioteconomia e da Organização do conhecimento (KO), é importante otimizar a troca de informações em domínios; portanto, os domínios precisam ter um certo nível de estabilidade e infraestrutura para serem bons candidatos para Análise de Domínio (Hjørland, 2017, p. 441, tradução nossa).

Portanto, a definição não limita domínios ao âmbito acadêmico, mas sim, amplia a concepção para indicar que é possível investigar um corpo de conhecimento compartilhado, ontológica e epistemologicamente, por seus sujeitos. É na ideia de domínio não científico que podemos expandir a abordagem para outros domínios focados em diferentes tipos de conhecimento, dentre os quais o conhecimento religioso e o conhecimento popular. Com base nisso, a Análise de Domínio tem na comunidade discursiva a unidade dos sujeitos que defendem e cultivam tais conhecimentos.

Para Almeida *et al.* (2007, p. 23), as comunidades discursivas são “aquelas formadas pelo pensamento, linguagem e conhecimento sincronizados de grupos sociais

distintos que fazem parte da sociedade moderna. São construções sociais constituídas por indivíduos e suas dimensões culturais, sociais e históricas”. É a partir desses grupos que se torna viável acompanhar as relações que estabelecem o consenso e o dissenso perante as bases ontológicas e epistemológicas. São através do pensamento e da linguagem que nos diversos ambientes institucionais, sejam reuniões, eventos, comunicações e publicações, a abordagem sócio cognitivista da AD acompanha o desenvolvimento de um domínio.

As trocas no ambiente científico, no trabalho e nas interações fixadas em registros e na produção de quaisquer produtos intelectuais, técnicos e culturais, determinam como os indivíduos dessas comunidades atuam. Dias (2015, p. 11) esclarece que essas comunidades “compõem-se de atores com pontos de vista distintos, estruturas de conhecimento individuais, predisposições, critérios de relevância subjetivos, estilos cognitivos particulares. Mas, se fazem presentes no jogo entre as estruturas de domínio e o conhecimento individual e na interação entre o nível individual e social”. Em resumo, as comunidades representam os sujeitos que constroem e interagem com a dinâmica do domínio.

Mai (2008) também trata comunidades discursivas como atividades humanas. Isso recorda que as diversas práticas sociais detêm a informação e o conhecimento enquanto subsídios seja para a sua formação intelectual e científica, seja para atividade profissional ou desenvolvimento social de quaisquer atividades. Hjørland (2004) a Análise de Domínio visualiza os usuários como indivíduos pertencentes a diferentes culturas, estruturas sociais e de compartilham domínio do conhecimento, além de serem qualificados como produtores, intermediários e usuários de informações, conectados ou parcialmente conectados em comunidades e que possuem comum idiomas, gêneros e outras práticas de comunicação.

O conceito de comunidade discursiva é apropriado para pesquisadores que adotam a “Visão Social” (Faigley, 1986 *apud* Swales, 1990), pois acreditam que o substrato de suas pesquisas deve lidar com agentes conectados à uma realidade, utilizando uma fala comum para defender e mediar seu corpo de conhecimento. Consequentemente, retoma as principais abordagens e métodos utilizados pelas ciências sociais, principalmente, aquelas dos pontos de vista antropológico e sociológico do conhecimento e da ciência.

Podemos associar que essa noção, em parte, reconhece as pessoas em suas atividades e segmentos da sociedade, das quais convivem com fatores sociais, políticos, econômicos, históricos e culturais. Além disso, carece de entender que a Análise de Domínio compartilha com o construtivismo social e interesse em perspectivas históricas acerca das condições sociais da produção do conhecimento (Hjørland, 2004a). Essas condições sociais podem ser identificadas nas comunidades discursivas através da investigação de documentos,

artefatos e quaisquer produtos produzidos ou imprimidos pela linguagem e o discurso que os membros de uma comunidade partilham.

Swales (1990), a partir dos seus estudos acerca de gêneros textuais e da teoria do discurso alocados na linguística e na sociolinguística, aponta que a comunidade discursiva é um termo que condiz com as 'perspectivas sociais' para seus diversos propósitos aplicados na pesquisa escrita. Entretanto, essa noção requer distinção de outros termos similares como é o caso de comunidades epistêmicas e de pensamento.

Situados nos estudos de domínio no âmbito da organização do conhecimento, Martínez-Ávila, Guimarães e Evangelista (2017) apontam que as comunidades epistêmicas podem ser assimiladas como redes de especialistas que compartilham noções de validade, trocam conhecimento e articulam-se como produtores de conhecimento com base em um conjunto de crenças. Pressupõe nessa definição que essas unidades epistêmicas formadas por pessoas com um certo controle, mediação e produção do conhecimento estão bem mais fixadas nos aspectos epistemológicos, denotando por exemplo, que a participação ativa de pesquisadores na formulação de teorias e outros construtos científicos ou de profissionais que convivem ativamente nas relações de gestão de conhecimento tácito e explícito.

Diante dessas linhas de exposição, entende que esses grupos dominam e utilizam o discurso como ferramenta de exposição e articulação do conhecimento, ou seja, é através do discurso que os indivíduos dominam seu conjunto de experiências, saberes e quaisquer outras expressões que caracterizam um domínio. Gregolin explica que “o DISCURSO é um suporte abstrato que sustenta os vários TEXTOS (concretos) que circulam em uma sociedade. Ele é responsável pela concretização, em termos de figuras e temas, das estruturas semi-narrativas (1995, p. 17, destaque da autora).

Alicerçado na teoria do discurso, Foucault (1996) infere que o discurso contribui para o desenvolvimento das disciplinas ao ser meio e expressão de novas perspectivas, a renovação de princípios e de formas analíticas em contextos científicos e sociais. Contudo, há mecanismos de controle dos discursos que impõem condições de funcionamento, imposição de regras e consequências que limitam o acesso. Assim, pessoas são adequar às condições, à ordem e aos gêneros¹³ do discurso. Recorrendo a Swales (1990), o gênero é uma categoria relevante na noção do discurso que caracteriza as comunidades.

Um gênero compreende uma classe de eventos comunicativos, cujos membros compartilham os mesmos propósitos comunicativos. Tais propósitos são reconhecidos pelos membros especialistas da comunidade discursiva de origem e, portanto,

¹³ Utilizamos aqui “gêneros” a partir da noção de Swales (1990) situada na Teoria do Discurso.

constituem o conjunto de razões (*rationale*) para o gênero. Essas razões moldam a estrutura esquemática do discurso e influenciam e impõem limites à escolha de conteúdo e de estilo (Swales, 1990, p. 58, tradução nossa).

A partir da definição de gênero, percebe-se como o discurso é estabelecido na comunidade e responde a uma série de questionamentos levantados na Análise de Domínio no tocante de suas funções e de que formas mediadas.

O discurso também pode servir de base para a manipulação e a distinção entre os indivíduos. As instituições, nesse caso, exercem seu domínio e controle sobre as formas discursivas, atingindo o sujeito pela imposição de regras, de condições e de limitações. Um exemplo básico, refere-se à construção discursiva desencadeada pelo uso do conceito “moral” no período da idade média.

Em uma das linhas da análise do discurso e considerando diversos tipos de discursos, dentre os quais políticos, religiosos e midiáticos, Van Dijk (2010) articula suas observações na relação entre o poder social e o discurso, partindo da análise geral desse poder e suas ramificações estruturais, seja em grupos, seja em instituições, ao declarar que há um domínio social e, em consequência, há respostas que representam a resistência e o surgimento do contrapoder. Ou seja, dentre os tipos de discursos correspondentes aos diversos grupos e ou comunidades existem aqueles que se contrapõem diante de outras coletividades e exercem formas de controle na sociedade.

5 QUESTÕES E ASPECTOS FUNDAMENTAIS ACERCA DOS LÉXICOS E DAS INFLUÊNCIAS ÉTNICAS

Para introduzir os estudos lexicais em contexto de diversidade étnica e vislumbrando os alicerces teóricos da Ciência da Informação e da Biblioteconomia, é necessário um resgate histórico e conceitual que permita contrapor e explicitar camadas de discursos densos e conflitantes, não somente no meio acadêmico e científico, mas em contextos sociais e antropológicos que apontam suas origens e aplicabilidades.

5.1 Introduzindo os conceitos de etnia, etnicidade, grupos étnicos, raça e nação

O primeiro ponto conflitante converge para as interpretações e usos dos conceitos de etnia, grupos étnicos, etnicidade, indígena, raça e racismo. A seguir o texto dissertativo aponta diferentes conceitos e opiniões de pesquisadores. Não é o objetivo utilizar um falso lugar de falar, mas sim descobrir e colaborar com futuros estudos de pesquisadores localizados no protagonismo e no seio desses grupos aqui mencionados. Ou seja, aqui apoia-se nos recentes e nos futuros estudos de pesquisadores que participam e se colocam nos seus lugares de fala¹⁴. O primeiro passo, aqui, revisar esses conceitos no campo científico da Antropologia e das Ciências Sociais conforme os relatos das teorias da etnicidade.

Nessa perspectiva, Poutignat e Streiff-Fenart (1998) introduzem um painel elementos e atores sociais, partindo para a complexidade do entendimento do que seria a identidade nacional e a integração de grupos de imigrantes. Grupos esses que podem ser confundidos com grupos étnicos ou levantar proposições equivocadas, fora da exploração investigativa e da localização do tema em um recorte temporal e epistêmico do qual o conduz às narrativas dos discursos científicos. Esses elementos e atores são peças-chave para compreender os modos de vida, alguns aspectos da categoria conceitual “cultura” e as questões políticas e econômicas das quais convergem para tal temática da etnicidade.

A partir da introdução e das primeiras premissas da teoria da etnicidade, nota-se uma inclinação de um discurso de percepção conceitual voltada para estudos anglo-saxões, dentre os quais os autores norte-americanos. Visualiza-se, ainda, que a noção de etnicidade está intimamente ligada ao “problema racial”, visto que para os estudiosos franceses destacam-se, primeiramente, com um de “problema imigração”. Ambas as inclinações devem ser localizadas

¹⁴ O conceito de “lugar de fala” explicado por Ribeiro (2019) como a referência que alguém profere seu discurso com base e pautado na sua realidade social e individual.

nos discursos fomentados desde o século XIX e, mais recentemente, em meados da década 1980, considerando que à luz das teorias instrumentalistas de que a etnicidade situa-se “[...] como uma reação às mudanças das estruturas institucionais e das relações de poder implicadas pela modernização” (Poutignat; Streiff-Fenart, 1998, p. 100), refutando as origens que a declaram ser um efeito de socialização entre grupos étnicos.

O termo etnicidade no contexto francês foi utilizado em 1981 durante uma mesa-redonda organizada por Françoise Morin em meio às atividades da Associação Francesa de Antropólogos e tornou-se um candidato a novo paradigma nas Ciências Sociais (Poutignat; Streiff-Fenart, 1998). O paradigma nomeado, aqui pelos autores, deflagra uma série de estudos que venham de algum modo justificar e reinterpretar o ser social nos acontecimentos históricos dos séculos XIX e XX, em contraponto aos contatos entre pessoas de diferentes nações e da transposição dos limites culturais - não mais atrelados às fronteiras geográficas.

Nota-se que a amplitude que envolvem os discursos da etnicidade existe em um plano caracterizado pelo espaço político, do qual requer a todo custo que o indivíduo ou o cidadão seja caracterizado e inserido em tal esfera política. Oliveira (2000, p. 178) aponta que essa noção de etnicidade remondada na antropologia “[...] nos induz a visualizar um panorama no qual se defrontam - melhor diria, confrontam-se - grupos étnicos no interior de um mesmo espaço social e político dominado apenas por um deles”.

Conceitualmente, no âmbito da literatura de língua inglesa, o termo tem sua origem durante a década de 1940 e suas primeiras implicações se referem a pertença de um outro grupo não anglo-americano, além de ser empregado como uma variável para estudar o comportamento dos indivíduos como foi o caso de Warner, primeiro autor a tratar do termo a partir do estudo *Yankee City Series* (Poutignat; Streiff-Fenart, 1998).

Outros exemplos históricos podem ser acompanhados nos estudos na década de 1970, dos quais têm-se a figura do “trabalhador imigrante” carregando os estigmas decorrentes da relação direta com as famílias francesas e seu convívio de adequação ou não adequação aos espaços, costumes e ou outras instituições. Para aprofundar tal painel social, o conceito e as noções de etnicidade utilizadas nas ciências sociais anglo-saxônicas são utilizados “[...] para descrever tais processos de atribuição categorial e de organização das relações sociais a partir de diferenças culturais presumidas essenciais” (Poutignat; Streiff-Fenart, 1998, p. 17). A problemática de entender uma noção de etnicidade amplamente divulgada na comunidade científica de língua inglesa leva a diferenciação e a atenção destinada à existência dos chamados “grupos étnicos”.

Entretanto foi o antropólogo norueguês, Fredrik Barth, que se dedicou à temática dos “grupos étnicos”, chegando às possíveis análises de tal termo e enveredando em estudos descritivos acerca da diversidade das comunidades. Como exemplo deste tema, visualiza-se estudos acerca da compreensão da origem de um grupo étnico e as características em relação às demais coletividades. A presença das acepções conceituais leva a menção de *ethncis* ou grupos étnicos para serem referidos àqueles núcleos designados e destacados por esses estudos. Nesse debate, uma das análises pode instruir:

[...] tal utilização da palavra “étnico” para designar contrastivamente e muitas vezes negativamente povos “outros” é congruente com as raízes etimológicas do termo etnicidade (*ethnikos*). No mundo grego, o termo *ethnos* fazia referência aos povos bárbaros ou aos povos gregos não organizados segundo o modelo da Cidade-Estado, ao passo que o termo latino *ethnicus* designava, na tradição eclesial do século XIV, os pagãos em oposição aos cristãos (Sollors, 1986 apud Poutignat; Streiff-Fenart, 1998, p. 23).

Tal indício histórico, pode denotar, em muitos estudos, a forma negativa como foi e ainda é a percepção geral desses grupos perante as demais instituições sociais e seus modos de vida e organização política e econômica. A igreja, nesse caso, esteve consolidada em propagar e justificar seu discurso de modo a diferenciar coletivamente as pessoas, primeiro com a oposição à outra religião e cultura, depois com a justificativa das características físicas e intelectuais de outros que não integravam o cristianismo e, em outros casos, o catolicismo.

Outro ponto de vista, alude a utilização desse termo para mencionar comunidades que sofreram e sofrem com a exploração dos países e nações colonialistas e detentoras de um imposto nível mais alto de desenvolvimento social e econômico. Além disso, o papel da segregação religiosa e a existência de instituições religiosas cristãs que alcançaram influência e poder sobre os demais seguimentos de natureza politeísta, dentre eles os africanos. Nesse sentido, Quijano aponta como a formação da América Latina colonial foi marcada pelo etnicismo. Essa entendida como a propagação e efetivação da institucionalização da etnicidade:

A formação do mundo colonial do capitalismo deu origem a uma estrutura de poder cujos elementos cruciais foram, especialmente na sua combinação, uma novidade histórica. Por um lado, a articulação de diversas relações de exploração e trabalho – escravatura, servidão, reciprocidade, salários, pequena produção comercial – em torno do capital e do seu mercado. Por outro lado, a produção de novas identidades históricas, “índio”, “negro”, “branco” e “mestiço”, posteriormente impostas como categorias básicas das relações de dominação e como fundamento de uma cultura de racismo e etnicismo (Quijano, 2014, p. 757, tradução nossa)

O etnicismo mencionado no texto implica na utilização da denominação étnica ou grupos étnicos para demarcar um processo histórico de estratificação social em função de um sistema colonialista, do qual apontava para dominar e, dentre outras ações, usufruir da força de trabalho dos povos, tanto originários das Américas, quanto os povos advindos do continente africano.

Recentemente, o termo etnicidade está presente nos estudos sobre imigração, racismo, nacionalismo e violência urbana. Dentre os estudos estão aqueles acerca dos diferentes grupos étnicos no contexto urbano, dos quais permitem analisar questões sociolinguísticas em países com diversidade cultural e linguística. Essa exploração do termo e sua noção não deve ser reduzida às questões específicas de determinados grupos étnicos, citados pelos primeiros teóricos de língua inglesa.

Alguns exemplos citados ilustram bem como os estudos de “[...] etnicidade por Glazer e Moynihan (1975) [em *Ethnicity: Theory and Experience*], eles se dedicaram a fenômeno tão diferente quanto os das majorias na China, das questões linguísticas na Índia, aos problemas de nacionalidade na União Soviética, ou a posição dos negros na sociedade americana” (Poutignat; Streiff-Fenart, 1998, p. 27). Citando assim, a existência de um apelo para encarar a etnicidade como temática mundial e se fazer necessária uma ampliação das discussões e dos diálogos em diferentes grupos epistemológicos e em todos os continentes e nações. Isso pode ser averiguado na leitura de teóricos contemporâneos da América Latina – Aníbal Quijano e Walter D. Mignolo – acerca da Decolonialidade e que notam a problemática cercada nas concepções universais, principalmente, aquelas que limitam e não condizem com estudos feitos por pesquisadores não europeus e norte-americanos.

A etnicidade é um fenômeno da modernidade por estar diretamente relacionado aos fatos e questões modernas, essencialmente contemporâneas, com enfoques e domínios próprios (Poutignat e Streiff-Fenart, 1998). Enfoques e domínios esses que precisam perpassar por processos de análises e aprofundamento crítico em virtude das novas epistemologias emergentes. Com o advento da Globalização e as mudanças no cenário político a partir do trânsito das culturas, Oliveira (2000) ressalta alguns dos problemas da etnicidade, além das dificuldades que uma comunidade de comunicação e argumentação intercultural se coloca em superar as novas normas para assegurar e regular um diálogo democrático.

Essas questões colocam esse conceito antropológico e social no centro de problemas atuais como aqueles dos direitos das comunidades indígenas frente às ocupações de terras por membros da elite branca, além do papel do estado democrático em garantir os direitos fundamentais descritos na Constituição brasileira. A questão de nacionalismo mencionada por

Poutignat e Streiff-Fenart (1998) também é confrontada nesse diálogo contemporâneo, no qual as etnias e a identidade de seus grupos perpassam por um julgamento e avaliação para qualificar se seus membros de fato pertencem a determinado grupo étnico-racial frente as políticas nacionalistas e governamentais. Daí a interpretação da etnicidade também se volta a qualificação daqueles oriundos de uma determinada etnia.

Quanto ao termo etnia levanta-se uma constatação essencial, principalmente, considerando as teorias da etnicidade: “trata-se ante de ver o modo como, desde sua criação no início do século XIX, a noção de etnia se encontra mesclada a outras noções conexas, as de povo, de raça ou de nação, como as quais mantém relações ambíguas cujo rastro encontramos nos debates contemporâneos” (Poutignat; Streiff-Fenart, 1998, p. 33). A primeira tentativa de desambiguação é a noção de etnia de Vacher de Lapouge introduzida nas ciências sociais:

Para este defensor inflamado da “escola selecionista”, a raça, definida como um conjunto dos indivíduos que possuem em comum um determinado tipo hereditário, é o fator fundamental da história. Zoologista antes de mais nada, como ele próprio se definiu, Vacher de Lapouge considera o homem não como um ser à parte, mas como um primata cuja característica de espécie é a de estar submetido mais à seleção social que à seleção natural. [...] Se Vacher de Lapouge inventa o vocábulo etnia, é, afirma eld para prevenir um ‘erro’ que consiste em confundir a raça – que ele identifica pela associação de característica morfológicas (altura, índice cefálico etc.) e qualidades psicológicas-, com um modo de agrupamento formado a partir de laços, intelectuais, como a cultura ou a língua. Tais grupos sociais (que ele define simultaneamente como ‘naturais e factícios’) não podem, segundo ele, confundir-se com a raça, e até mesmo lhe são ‘mais ou menos opostos’, dado que se trata de agrupamentos que resultam da reunião de elementos de raças distintas que se encontram submissos, sob o efeito de acontecimentos históricos, a instituições, a uma organização política, a costumes ou ideias comuns (Poutignat; Streiff-Fenart, 1998, p. 33-34).

Estas proposições de Vacher se concentram, primeiramente, em uma noção de raça que se opõe às recentes reflexões sobre as consequências negativas do preconceito e da segregação diante da identidade equivocada e baseada, exclusivamente, nos traços físicos e características genéticas, dentre eles cor da pele, textura do cabelo e cor dos olhos. Contestando, esta noção, Munanga (2004) esclarece que o conceito de raça, centrado na classificação de traços físicos e na justificativa biológica e genética, é pautado em uma teoria pseudocientífica articulada para fundamentar um discurso ideológico destinado às relações de poder e dominação.

Assim com a problemática da justificativa classificatória e de segregação destacada por Munanga (2004), Quijano (1999) se apoia na ideia de que a raça é uma perspectiva histórica de classificação humana e social que organiza e estabelece relações de dominação e instrumento de poder das nações colonizadoras. Essa ideia está no cerne da questão de poder e de

configuração das identidades culturais mencionadas por Quijano (1999; 2014) no contexto da América Latina. O que atesta uma fragilidade nos argumentos acima citados que justificam a classificação dos traços físicos humanos para incumbir uma teoria étnica. Essa perspectiva teórica é baseada em identidades justificadas exclusivamente por características biológicas em situações históricas e geográficas, das quais estão, em muitos casos, distantes da realidade retratada por teóricos e pesquisadores que não fazem parte ou não se identificam com tais identidades colocadas em análise.

Outra proposição de Vacher acerca da etnia fomenta a busca pela definição de parâmetros que possam encontrar, no seio das relações sociais, fatos e o modo como a sociedade é organizada – pessoas que compartilham os mesmos costumes e indícios de uma existência em comum. Essa busca esteve carregada pela relação dos traços biológicos (do conceito de raça) e dos traços intelectuais, uma vez que, grupos e instituições sociais associam a cor da pele branca aos melhores níveis de formação intelectual e desenvolvimento sociocultural.

Esse anseio pela compreensão das etnias, novamente é problematizado por Quijano (1999) no relato de outra característica do poder configurado para dominar e destituir as sociedades e culturas originárias, ou seja, a perspectiva de definir as populações colonizadas à categoria “índios”. Nesse sentido, essas populações, assim como aquelas de origens africanas, sofrem com a redução imposta por uma ideia de etnia abrangente e homogênea, desmerecendo os aspectos conceituais que consideram as etnias diante da heterogeneidade e no tocante da riqueza e da diversidade cultural.

A discussão conceitual da etnicidade e da raça entorno da colonialidade do poder em Quijano (1999), também pode ser entendida na articulação do capitalismo histórico ambientado na geopolítica do conhecimento e as desigualdades coloniais quando o autor cita Wallerstein (1983 apud Mignolo, 2020) para indicar as consequências fundamentais da “etnicização da força do trabalho” na constituição do sistema colonial moderno e do racismo, são elas: a atribuição da força de trabalho de “pele negra a escravatura”; a etnicização responsável por viabilizar um mecanismo de enquadramento agregados etnicamente definidos (da mesma família) e não à custa dos empregadores ou do estado; e a etnicização enquanto racismo institucional e pilar do capitalismo histórico.

Depois de elucidar a conceituação apresentada por Vanche, Poutignat e Streiff-Fenart (1998, p. 35) declaram essa noção como sendo uma definição negativa e que coloca em oposição os laços biológicos e laços intelectuais. Apesar dessa dualidade existir, é com base nela que a reivindicação pela adequação da aplicabilidade dos conceitos e sua natureza preconceituosa tende a dispensar e reconstruir teorias e proposições opressoras, assim

destacando diferenças complexas entre as comunidades analisadas pelos estudos antropológicos e sociológicos. Nessa ocasião, Raffestin (1993) é direto quando afirma que as diferenças raciais e étnicas são constituídas por fator político de modo virtual ou concreto e que essas diferenças concebidas na História são delineadas por uma onda contínua, por si mesma uma vontade de poder e apoiada em preconceitos.

Poutignat e Streiff-Fenart (1998) encontram uma distinção com base nas leituras de Renan e Weber, destacando, também, a desambiguação provocada pelo uso do termo nação:

O que distingue a pertença racial da pertença étnica é que a primeira é “realmente” fundada na comunidade de origem, ao passo que o que funda o grupo étnico é a crença subjetiva na comunidade de origem. Quanto à nação, ela é, como o grupo étnico, baseada na crença da vida em comum, mas se distingue deste último pela paixão ligada à reivindicação de um poderio político (Poutignat; Streiff-Fenart, 1998, p. 37).

A tríade conceitual deste trecho evoca as relações que partem: do indivíduo para o seu grupo de convívio e unidade do qual ele está vinculado; do grupo para os demais grupos que compartilham de um espaço geográfico politicamente institucionalizado; e, por último, do somatório de aspectos desses grupos que compõem o processo de elucidação daquilo que se entende por nação. Ou seja, com base nesse trecho podemos ainda recorrer ao conceito de identidade, e mais precisamente, a identidade étnica.

Analisando a definição de etnia, podemos esclarecer alguns aspectos destacados na noção de identidade étnica defendida por Poutignat e Streiff-Fenart (1998) como sendo essa uma construção dada a partir da diferença e o que atrai um indivíduo ao seu grupo é indissociável da questão da repulsa daqueles tidos como diferentes e estrangeiros, ou seja, “[...]esta ideia implica que não é o isolamento que cria a consciência de pertença, mas, ao contrário, a comunicação das diferenças das quais os indivíduos se apropriam para estabelecer fronteiras étnicas” (*Ibid.*, p. 40).

Nesse sentido, entende-se que estamos diante de duas descobertas acerca da investigação empírica do caráter das fronteiras étnicas. Além disso, podemos desmistificar a visão simplista que defende os isolamentos sociais e geográficos como fatores cruciais na manutenção da diversidade cultural. Pois primeiramente, as categorias étnicas, aquelas tidas como objeto empírico, não dependem da falta de mobilidade, de contato ou de informação, mas acendem de forma efetiva para os processos de exclusão e de incorporação, mesmo que diante das mudanças de participação e de pertencimento dos indivíduos (Barth, 2000).

Isso abre espaço para entender o papel da linguagem e dos costumes em detrimento de tais fronteiras, além desses serem intrínsecos à natureza de cada grupo étnico. Em segundo

lugar, Barth infere que as distinções étnicas não dependem da ausência de interação e aceitação sociais, do contrário, essas distinções são, de maneira frequente, “[...] a própria base sobre a qual sistemas sociais abrangentes são construídos. A interação dentro desses sistemas não leva à sua destruição pela mudança e pela aculturação: as diferenças culturais podem persistir apesar do contato interétnico e da interdependência entre etnias” (Barth, 2000, p. 26).

Barth (2000) aponta que para a literatura antropológica o grupo étnico, pode ser compreendido de tal modo que denota uma população como sendo aquilo que: em maioria se autoperpetua de forma biológica; compartilha de valores culturais fundamentais; compartilham de um campo de comunicação e interação; e tem um conjunto de membros que se distingue e é identificado por outro. Essas quatro características predicativas do conceito de grupo étnico são problematizadas pelo autor:

Mesmo assim, na forma modificada apresentada acima, não deixa de ser uma definição que se aproxima suficientemente de muitas situações etnográficas empíricas, ao menos tal como elas apareceram e foram relatadas, de modo que esse significado continua a ser útil para os objetivos da maioria dos antropólogos. Minha discordância não diz respeito tanto ao conteúdo substantivo dessas características, ainda que eu pretenda demonstrar que temos muito a ganhar com uma certa mudança de ênfase; minha principal objeção é que, tal como está formulada, essa definição nos impede de compreender o fenômeno dos grupos étnicos e seu lugar na sociedade e na cultura humanas. Isso porque ela evita as questões mais fundamentais: ao tentar oferecer um modelo típico-ideal de uma forma empírica encontrada recorrentemente, essa formulação traz implícita uma visão preconcebida de quais são os fatores significativos para a gênese, a estrutura e a função de tais grupos. O principal problema desta visão é o seu pressuposto de que a manutenção das fronteiras não é problemática, e que isto se dá como consequência do isolamento que as características arroladas implicam; diferenças racial e cultural, separação social, barreiras linguísticas, inimizade espontânea e organizada (Barth, 2000, p. 27-28).

Ou seja, há um modelo empírico que desconsidera as causas e os efeitos de tais negatividades expressadas em tal entendimento conceitual e no que ele pode influenciar no processo metodológico que visa descobrir os grupos étnicos e suas características.

Para examinar os contatos interétnicos, Oliveira (1976) busca no conceito de identidade étnica uma forma de assimilar a noção de identificação suscitada a partir da proposta de grupos étnicos. Esse autor recorda que o ponto central de observação está no aspecto cultural, que para Barth (2000) é, a priori, a característica primária no processo de identificação desses grupos. Através dos “contatos culturais”

O que se tratou de entender então, porém, foram as relações de identidade-status no âmbito de uma cultura determinada (e guardada sua inteireza, ao menos para fins analíticos), e não de se penetrar no interior de um outro sistema de relações, sincrético por natureza, tal como o sistema interétnico - este sima base de inteligibilidade da identificação étnica. A "cultura do contato", entendida principalmente como um

sistema de valores altamente' dinâmico, portanto susceptível de fornecer o *rationale* das "flutuações" da identidade étnica (ou, em outros termos, a lógica da manipulação dessa identidade), poderá permitir a elaboração de uma tipologia capaz de conter diferentes "culturas do contato" e de conformidade com a maior ou menor distância e "oposição" das culturas em conjunto, da maior ou menor tensão e conflito entre os grupos étnicos em contato. Nesse sentido, essa "cultura do contato" pode: ser mais do que um sistema de valores, sendo o conjunto de representação (em que se incluem também os valores) que um grupo étnico faz da situação de contato em que está inserido e nos termos da qual classifica (identifica) a si próprio e aos outros (Oliveira, 1976, p. 23).

Nesse sentido, para captar os traços e indícios culturais, vale-se da identificação étnica a partir da identidade (individual e social) e das modalidades de identificação em contextos intertribais e nos confrontos das relações sociais. Em parcela significativa, esses contextos estão concentrados em situações das quais os indivíduos se assemelham – grupos étnicos da mesma localização ou com trações comuns – e no cenário de embate racial travado com os brancos, segundo processos de dominação e estabelecimento de relações de poder e trabalho.

Essa compreensão complexa de identidade refaz uma discussão que supera as questões culturais e integra o limiar entre o estado psíquico e psicossocial com o cenário de confronto social, não somente étnico, mas racial (Oliveira, 1976). Alude, ainda, aos confrontos manifestados conforme os ditames e a organização das estruturas de classes, ambos contrastados pelas influências ideológicas e representações sociais que moldam o indivíduo e suas relações.

Antes de existir uma teoria da etnicidade, historicamente, a etimologia e o modo de poder fixado no conceito de raça e etnia denotam significados opressores, tidos como força que rege as formas de preconceito e dominação. É nessa primeira premissa que “no latim medieval, o conceito de raça passou a designar a descendência, a linhagem, ou seja, um grupo de pessoas que têm um ancestral comum e que, *ipso facto*, possuem algumas características físicas em comum” (Munanga, 2004, p. 17).

Abaixo, destacaram-se alguns tópicos e trechos importantes identificados nas teorias e na discussão conceitual aqui destacada. Poutignat e Streiff-Fenart (1998) no texto “As evidências do empirismo e as ingenuidades do etnólogo: Leach, Moerman, Barth”, os autores apontam três principais ingenuidades:

A primeira ingenuidade é a de se acreditar que se pode definir uma unidade étnica (quaisquer que sejam os critérios utilizados para defini-la) por uma lista de traços. Em sua resposta a Naroll, Leach denuncia o absurdo das descrições taxonômicas (a produção de “contra-senso tabulado”). Barth e seus colaboradores demonstram ser impossível encontrar um conjunto total de traços culturais que permitam a distinção entre um grupo e outro, e que a variação cultural não permite por si própria abranger

o traçado dos limites étnicos. A realidade a qual se dobra o antropólogo é que identidades distintas podem ser mantidas na ausência de traços culturais comuns comprovados e que, ao inverso, uma teoria indígena da diversidade étnica pode existir apesar da homogeneidade cultural constatada pelo observador. [...] A questão, a partir daí, não é mais estudar a maneira pela qual os traços culturais estão distribuídos, mas a maneira como a diversidade étnica é socialmente articulada (Poutignat; Streiff-Fenart, 1998, p. 61-62).

As reflexões dos autores assinalam mais para a importância de a diversidade étnica existir do que mesmo uma tentativa objetiva de o pesquisador tentar elencar características que distinguem um grupo étnico em detrimento de outros grupos, étnicos ou não. Ou seja, o ato de explicar o que define ou não um grupo étnico, com base no modo ou metodologia descritiva, torna-se problemático, na medida em que o essencial está em como o ser “étnico” se manifesta socialmente e como ele é percebido e destituído de seus direitos perante uma sociedade ou uma instituição social.

Segunda ingenuidade: a de acreditar que o isolamento geográfico e social esteja na base da diversidade étnica. A investigação empírica, diz Barth, produz duas descobertas que demonstram a inadequação desta visão das coisas: primeiramente, as fronteiras étnicas persistem apesar do fluxo de pessoas que as atravessam. Em segundo lugar, relações frequentemente de uma importância vital são mantidas através das fronteiras. A interpenetração e a interdependência entre os grupos não devem ser vistas como dispersões das identidades étnicas, mas como as condições de sua perpetuação. Moerman (1965, 1968), por sua vez, mostra que, no caso dos luses da Tailândia, não somente a diversidade étnica persiste apesar do contato, mas que ela só adquire sentido no seio de uma área expandida de homogeneidade cultural e linguística (Poutignat; Streiff-Fenart, 1998, p. 62).

A menção dessa ingenuidade corrobora com o que Barth (2000) afirma sobre as fronteiras sociais serem aquelas das quais precisamos nos concentrar, pois quando observamos as fronteiras territoriais, pode-se entender que a identidade de um grupo se mantém mesmo que seus indivíduos estejam interagindo fora de seu território e com outros grupos.

Poutignat e Streiff-Fenart (1998, p. 63) aponta que a terceira ingenuidade é fixada em “[...] acreditar que um rótulo étnico = um modo de vida = um grupo real de pessoas, enquanto é precisamente a análise da relação problemática entre esses três elementos que o etnólogo deveria aceitar como tarefa”. Esse rótulo prescreve um equívoco e deve ser analisado com cautela, tendo em vista que a confusão terminológica que tornam equivalentes ou semelhantes conceitos como raça, etnia e nação, fazendo com que suas interpretações sejam imprecisas e conflituosas.

Representante dos estudos antropológicos africanos e sul-americanos, Kabengele Munanga contribui conceitualmente e historicamente no debate dessas confusões terminológicas e epistemológicas. Segundo o autor, o conceito de raça advém da palavra

italiana *razza*, essa por sua vez, tem origem do latim de *ratio*, do qual significa sorte, categoria, espécie, e foi utilizada primeiramente na Zoologia e na Botânica para classificar as espécies de animais e vegetais (Munanga, 2004). A partir dessa constatação histórica, a etimologia impôs, desde o seu uso, a designação de modelos de classificação de seres, como plantas ou outros seres vivos, sempre com o intuito de emitir um modelo de identificação entre semelhança e diferenças até então tidas como características físicas, genéticas e biológicas.

5.2 Aspectos gerais dos estudos dos léxicos

O contato com o léxico é o primeiro e principal passo para conhecer uma determinada comunidade e seus costumes. A língua molda o pensamento, a percepção e a leitura de mundo, viabilizando as experiências culturais (Caldas; Barros; Silva, 2016). Diferente de sistemas fechados (fonologia, morfologia e sintaxe) o léxico pode ser considerado como o único domínio da língua tido como um sistema aberto (Biderman, 2001). A abertura se efetiva na liberdade com que seus falantes utilizam e remodelam constantemente as unidades lexicais, considerando que no cotidiano a língua é dinâmica e se adapta para os diferentes propósitos da comunicação entre os sujeitos.

As fronteiras dos modernos estudos linguísticos acerca dos léxicos se expandem para a Lexicologia, Lexicografia, Terminologia, e Terminografia, chegando aos outros ramos da Linguística, como a Dialectologia, a Sociolinguística e a Etnolinguística, pois as marcas dialetais, sociais e culturais dos falantes estão refletidas na língua, de um modo geral, e nas particularidades do léxico (Aragão, 2010). Desse modo, o percurso de compreensão dos léxicos inicia-se com a atenção dada para a língua e suas particularidades em meio as relações humanas.

Para Saussure (1977) a língua é um produto social, fenômeno psíquico, processo fisiológico e processo físico, reiterando os fatos da consciência (conceitos) associados às representações dos signos linguísticos ou imagens acústicas. Os signos linguísticos atuam nesse processo como unidades elementares e intrínsecas da língua. O autor esclarece que a língua pode ser compreendida como sistema composto por signos socializados e “é, ao mesmo tempo, um produto social da faculdade de linguagem e um conjunto de convenções necessárias, adotadas pelo corpo social para permitir o exercício dessa faculdade nos indivíduos” (*Ibid.*, p. 17).

É necessário nesse interim, recordar que a definição de língua é estreitamente ligada à de linguagem. Essa última eleva a condição das características específicas e usos das línguas em diferentes contextos, propriedades físicas e aspectos culturais próprios, ou seja, a linguagem

é vista como a forma e o modo como passamos a comunicar nosso pensamento. Melo e Bräscher (2011, p. 22) explicam que “a linguagem é o que dá forma às mensagens”. Essa ideia é aquela que vai nos ajudar a entender que uma linguagem de especialidade se constitui, além de normas e convenções como as línguas, de qualidades e propriedades ancoradas em discursos científicos e técnicos, dentre os quais: seguir os padrões de comunicação em espaços e ambientes institucionais; produzir jargões técnicos; e serem disseminadas em documentos e sistemas preestabelecidos. Diferentemente, as línguas detêm uma característica mais primitiva e complexa, aquela dada pelo ato da fala e seu uso amplamente difundido e estabelecido para comunicação de pessoas de uma determinada região geográfica, países e comunidades étnicas.

A alusão da complexidade das línguas, principalmente, o destaque para as línguas indígenas, pode ser conferida no trecho abaixo:

As línguas do mundo são classificadas em famílias segundo o critério genético. De acordo com esse critério, uma família linguística é um grupo de línguas para as quais se formula a hipótese de que têm uma origem comum, no sentido de que todas as línguas da família são manifestações diversas, alteradas no correr do tempo, de uma só língua anterior. As línguas românicas ou neolatinas - Português; Espanhol, Catalão, Francês, Romana, Italiano, Romeno - constituem uma família, cujos membros derivam de uma língua ancestral bem conhecida historicamente - o Latim. Para a maioria das famílias linguísticas, porém, as línguas ancestrais são pré-históricas, não se tendo delas nenhuma documentação. O conhecimento dessas línguas (ou de, pelo menos, certas características delas) é obtido mediante estudos histórico-comparativos que, partir da descoberta de correspondências regulares (de sons, de palavras, de formas gramaticais) entre duas ou mais línguas, formulam hipóteses sobre as propriedades que devia ter uma língua ancestral para permitir (e explicar) a derivação diferenciada das línguas atuais (Rodrigues, 1986, p. 29).

A configuração das línguas, nesse sentido, reverbera as heranças e raízes de um povo, conforme os percursos históricos e culturais que detêm atenção dos estudos históricos-comparativos, demonstrando as transformações e mudanças nos léxicos, como por exemplo os fenômenos de aglutinação verificados na formação e derivação morfológica de palavras. Colaborando com as reflexões, Corrêa-da-Silva (2013) sugere a possibilidade de conhecer e descrever as informações do mundo expressas por uma comunidade atual ou pré-histórica, mormente através da análise do léxico compartilhado pelos seus membros. Ou seja, uma tentativa de conhecer uma língua indígena corresponde a regatar e descobrir a comunidade de seus falantes.

Ao encontro da definição acima de Saussure, Lyons (2011), durante um estudo introdutório da Linguística, discute e analisa as definições de “língua (gem)” de alguns teóricos, dentre eles Chomsky, e chega ao ponto em comum: de que as línguas são sistemas de símbolos projetados para a comunicação. É nessa condição do ato de comunicar que essas podem ser

notadas enquanto produto social, e para além disso, os fenômenos linguísticos se mostram intimamente ligados aos fenômenos sociais.

Inicialmente, a língua assume um ponto de convergência entre diversos outros sistemas, como a semântica que a partir do contexto da fala e do registro dos signos linguísticos busca estabelecer os significados e interpretações. Nesse sentido, o sistema linguístico é composto “[...] de subsistemas (fonológico, morfológicos, sintático e semântico) que possuem cada um, por sua vez, partes menos que denominamos unidades elementares. Assim, o estudo da Fonologia faz-se a partir de fonemas; a Morfologia, de morfemas; a Sintaxe, sintagmas e a Semântica, de semas” (Melo; Bräscher; 2011, p. 27).

Verificando essa sistemática, Biderman (2001, p. 23) explica que o léxico gerado “[...] se processou e se processa através de atos sucessivos de cognição da realidade e de categorização da experiência, cristalizada em signos linguísticos: palavras”. Nessa concepção, a efetividade do léxico concentra-se no processo de nomeação e na forma de registrar as entidades e objetos. O registro compete, aqui, ao mecanismo de nomear, classificar e atribuir características e propriedades para todos os seres e objetos concretos ou abstratos, transformando, assim, em signos linguísticos representados durante os processos e meios de comunicação.

Contudo, conceitualmente, Schmitz (2015, p. 10) recorre a etimologia e verifica que o vocabulário do léxico se refere ao uso das palavras em diversas situações sociais, do qual “[...]no grego, *leikos* significa “sobre palavras” e origina-se de *lexis* (“palavra”) relacionado ao verbo *legein* (“falar”)”. Assim, é naturalmente que o próprio entendimento do léxico perpassa por questões conflituosas que coincidem em modelos e modos pelos quais palavras são redistribuídas e justificadas nos processos de comunicação.

A palavra aqui assume a unidade que assegura o registro e a forma, dos quais dependem integralmente da língua e da linguagem que buscam apresentar. Por sua vez, são apresentadas em diferentes símbolos gráficos, sejam alfabetos (representações de unidades elementares de fonemas e morfemas) e ideogramas (símbolos e signos responsável por expressar a ideia de objetos, ações e entidades).

Diante desses elementos, se faz necessário a distinção entre palavras e vocábulos, logo nas implicações da distinção entre léxico e vocabulário. Melo e Bräscher (2011) esclarecem que as palavras se constituem de forma livre na língua e na teoria linguística seu termo técnico é “lexemas”. Nessa acepção, os lexemas compõem o conjunto do qual chamamos de léxico. Enquanto as palavras focam em evidenciar denominações na língua, os vocábulos, por outro, lado implicam em outros aspectos formais de ligações e de referência, essenciais no

entendimento sintático e semântico. Ou seja, os “vocábulo – em contraste com as palavras – não traduzem ideias (possuem apenas significado gramatical) e são considerados “instrumentos gramaticais”, pois servem para estabelecer *relações* entre as palavras [...]” (Ibid. p. 83). Em alguns casos, um vocábulo pode ser considerado palavra, em outros não. Contudo, toda palavra é considerada vocábulo, pois são instrumentos gramaticais.

A associação de palavras aos conceitos é tida como estratégia no processo de estabelecimento classificatório, pois funcionam como pontos referenciais e auxiliares na identificação e associação dentro do universo de uma língua natural. Já “os símbolos, ou signos linguísticos, se reportam ao universo referencial” (Biderman, 2001, p. 14). A língua nesse sentido permite a fluidez com que o nosso pensamento é transmitido para o âmbito social.

O léxico é composto por todas as palavras de uma língua. Cada palavra é estabelecida conforme os subsistemas linguísticos mencionado acima (fonológicos, morfológicos, sintáticos e semânticos). Por exemplo, os fonemas constituem as unidades sonoras durante o processo de emissão dos sons das palavras. Melo e Bräscher (2011) explicam que, no caso dos léxicos, o léxico de uma língua possui o subconjunto dos campos lexicais que investigam as semelhanças de continuidade, sinonímia, hiponímia, entre outras noções de relações, das quais são responsáveis por estabelecer as entradas das unidades lexicais, chamada de *lexemas*, presentes nos dicionários e em quaisquer outros registros escritos. Essas relações e a compreensão dessas unidades lexicais estão inseridas no processo de síntese na representação temática de documentos.

Borba (2007, p. 57) explica que os léxicos possuem propriedades sintáticas, semânticas e pragmáticas, tendo em vista que a investigação “[...] relaciona-se com o alcance dos signos lingüísticos, ou seja, a dimensão sintática [= relação signo – signo], a dimensão semântica [= relação dos signos com os objetos a que se aplicam] e dimensão pragmática [= relação entre os signos e seus usuários]”. Todas essas dimensões retomam as relações entre conceitos/termos estruturadas em tesouros, taxonomias, classificações, listas de cabeçalhos de assunto e ontologias. As implicações na interpretação de significados e uso de um determinado *lexemas* requer o conhecimento dessas dimensões, pois cada uma examina aspectos funcionais não somente atrelados ações de comunicação, mas a própria replicação em documentos e até mesmo na aceitação de mudanças pela comunidade falante.

A partir daí nota-se as nuances da metáfora “campo minado” descrita por Schmitz (2015) em alusão às instâncias de crítica que verificam o uso e a pronúncia dos itens lexicais pelos indivíduos. Ou seja, o léxico atravessa a ordem das condições e das conformidades a serem seguidas, assim como em outros fatos e fenômenos sociais que requer dos indivíduos

submissão aos protocolos e padrões a serem seguidos para, então, serem aceitos nessas situações sociais.

Parte da convivência em grupo é demarcada pela competência de entender e utilizar o léxico durante o aprendizado de uma língua e as trajetórias de formação social e intelectual. O reconhecimento de uma cultura, por exemplo, pode ser adquirido conforme o estudo de seu léxico, seja traduzido ou estudado a partir de documentos. Ou seja, “[...] as relações construídas pelos contatos culturais se efetivam e ganham referenciais por mediação do léxico, sobre o qual incidem os percalços dos processos interculturais e dos atos tradutórios” (Caldas; Barros; Silva, 2016, p. 121).

Segundo Biderman (2001), o léxico de uma língua natural pode ser considerado como patrimônio vocabular de uma comunidade linguística e essa língua tem o seu universo conceitual como um sistema ordenado e estruturado de categorias léxico-gramaticais. Isso implica na denominação de língua natural entendida como uma língua amplamente falada e articulada como principal característica comunicativa de uma comunidade linguística, sendo diferenciadas das línguas artificiais elaboradas com propósitos e interesses preestabelecidos.

Nesse sentido, podemos compreender que a definição de léxico a constituição de uma língua perante o seu estabelecimento representacional das categorias léxico-gramaticais portada de apresentação falada e escrita. Resumidamente, Krieger (2017) entende que o léxico está incluído em um conjunto mais amplo do qual chamamos de língua ou sistema, sendo as Lexicografia e a Lexicologia tidas como ciências dos léxicos, juntamente com a Terminologia que estuda os léxicos em contextos de especialidades e comunicação científicas.

A Lexicografia é a ciência dos dicionários, historicamente, iniciada de modo efetivo nos séculos XVI e XVII com a elaboração dos primeiros dicionários monolíngues e bilíngues, esses últimos para indicar palavras em latim e em outras línguas modernas (Biderman, 2001). O que consolidou a análise dos significados das palavras como principal objetivo da Lexicografia. Essa ciência esteve voltada para a prática e a técnica de elaboração de instrumentos que orientem o uso e as formalidades dos léxicos, apresentando, dentre outros pontos, a categorização gramatical das palavras de uma língua e suas definições que englobam a questão pragmática.

O dicionário de léxico apresenta uma série de informações etimológicas, gramaticais, entre outras, diferente do dicionário terminológico cujo intuito está na definição contextualizada de um determinado texto. As definições lexicográficas e as enciclopédias, descritas nos dicionários e nas enciclopédias relativas às linguagens e às línguas naturais, tem em comum com as definições terminológicas o estudo de unidades lexicais. Isso ocorre em

virtude de os termos técnico-científicos também integrarem elementos paradigmáticos e do uso dos léxicos no cotidiano de todo e qualquer falante dentro e fora dos seus espaços e discursos de especialidades.

Krieger e Finatto (2004) explica que a definição dicionarista, diferente da definição terminológica, engloba uma dimensão geral e explora uma quantidade de informações mais abrangentes no léxico. Em outros termos, O dicionário de léxico apresenta uma série de informações etimológicas, gramaticais, entre outras distintas daquelas contidas no dicionário terminológico cujo intuito está na definição contextualizada de um determinado texto. Diante da funcionalidade desse instrumento da lexicografia, Biderman (2001) considera o dicionário um objeto cultural de importância para as sociedades contemporâneas, do qual tem o intuito de descrever vocabulários da língua e definir os signos lexicais e seus respectivos conceitos.

Quanto a Lexicologia é uma “[...] ciência antiga, tem como objetos básicos de estudo e análise da palavra, da categorização lexical e da estruturação do léxico” (Biderman, 2001, p. 16). Diante dessas preocupações, os dicionários tornaram-se instrumento de registro e produto lexicológico. Embora, a preocupação dessa ciência incide diretamente nas categorias léxico-gramaticais e em padrões linguísticos que indicam o compartilhamento comum de uma língua. Sua investigação aprofundada remete-se às línguas naturais, dentre os tópicos de sua sistemática, seus padrões gramaticais, suas variações e procedimentos de tradução.

A questão do léxico na Lexicologia permite reconhecer a pluralidade das línguas, as questões voltadas para os idiomas e os dialetos, confrontando-se com as línguas de povos originários e étnicos. Por isso, diante das intenções e prerrogativas dessas investigações lexicais “justifica-se a diversidade de campos gramaticais, linguísticos e discursivos que se voltam ao léxico ou com ele se interconectam, razão por que ele se torna um ponto de cruzamento de diversos estudos linguísticos” (Krieger, 2017, p. 152).

A abrangência desses estudos esteve presente em documentos registram e evidenciam o repertório lexical, dentre os quais os glossários e os atlas. Cada um desses exercem um propósito e integram parte dos relatos linguísticos, geolinguísticas, sociolinguísticos. Aguilera e Silva (2022) destacam que os atlas linguísticos resgatam e registram a distribuição das variantes linguísticas, das quais carregam marcas étnicas-raciais e histórico-culturais. Desse modo, os atlas, os dicionários e os demais documentos linguísticos, fornecem um mapeamento viável de identificação do léxico em relação à própria construção histórica da língua, sua expansão e amplitude geográfica, assim como as influências que estiveram reintegrando e remodelando seu uso.

No caso da região amazônica, Caldas, Barros e Silva (2016) explicam que a presença dos léxicos das línguas indígenas pode ser repertoriada nos hábitos, na indumentária e nos demais artefatos criados e presentes no cotidiano. Nesse contexto, as autoras introduzem a discussão intercultural e tradutória na compreensão de vocábulos de plantas comuns presente na alimentação dos habitantes da região do Caeté, como o caso da madicoera (ou manicoera).

Nesta reflexão, especialmente, estudos sobre léxico de plantas nos orientam para a presente discussão intercultural e tradutória a partir da apreensão vocabular sobre mandicoera ou manicoera presentes no uso dos falantes da região do Caeté. Do universo lexical presente no uso cotidiano, a escolha do léxico mani ou mandinos pareceu propícia, tanto diante da investigação lexicológica e, mais especificamente terminológica, quanto para simbologia de seus derivados e correlacionados como a farinha de mandioca, bem como pela tentativa de compreensão por parte de usuários a respeito de se usar uma expressão e não outra. De outro modo, também a distância de significado que se opera diante das línguas em contato, tanto para as gerações mais jovens, quanto para os imigrantes na região suscita que trabalhos de descrição linguística retomem a temática (Caldas; Barros; Silva, 2016, p.127-128).

Essa temática evidenciada por Calda, Barros e Silva é um exemplo que aponta para dois fatos importantes nos estudos dos léxicos de povos originários do Brasil. Primeiramente, a questão de parte dos nossos costumes culinários e dos nossos saberes – acerca das plantas, elementos e seres originários dos biomas brasileiros – advêm das tradições indígenas presentes, em muitos casos, no léxico do português falado no Brasil. O segundo: é através das investigações dos léxicos, em consonância com o reconhecimento das origens culturais, que se tornam viáveis o resgate da memória e da manutenção de parte dessas línguas diante dos conflitos e problemas recorrentes na sobrevivência desses povos.

Diante dessa temática, recorreu-se as particularidades das unidades lexicais situadas no universo dos discursos etno-literários. Para averiguar esse universo, Barbosa (2009, p.139) explica que o estudo da etno-terminologia é uma possibilidade, pois parte dos “[...] discursos etno-literários: literatura oral, literatura popular, literatura de cordel, fábulas, lendas, mitos, folclore e discursos das linguagens especiais com baixo grau de tecnicidade e de cientificidade”.

As qualidades absorvidas das línguas de especialidades e da linguagem literária, fazem dos léxicos objetos inteiramente ligados e dependentes do valor semântico social e dos documentos descritos como históricos da cultura. O ponto central nesta consolidação disciplinar implica na compreensão dos léxicos a partir dos discursos, como os etno-literários. Estes presentes nos documentos, assumem a proposta de, mais do que registrar, carregar as concepções de mundo e os traços culturais e sociais de uma determinada comunidade, considerando a dinâmica da língua falada e escrita.

O universo de discurso, para Barbosa (2006), se estabelece e é capaz de se renovar incessantemente em rede de relações intertextuais (textos manifestados e enunciados) e de relações interdiscursivas (processos discursivos e sua produção). Complementando, Pais (1984) define um universo de discurso como um conjunto não finito (tendem *ad infinitum*) de todos os discursos expressados, dos quais apresentam características comuns são suscetíveis; às normas frásticas, lexicais, semânticos-sintáticas, fonético-fonológicas); ou às normas transfrásticas, narrativas, discursivas; ou aquelas, dependentes persuasão/interpretação, manipulação/contramanipulação; às formulações mediadas pelas relações intersubjetivas, espaciais e temporais de enunciação e enunciado.

Assim, a investigação lexical requer o reconhecimento, também, dos espaços e das características relativas aos discursos nesses contextos de diversidade linguística e cultural, levando em consideração que o léxico de qualquer natureza étnica pode representar diversos discursos e ser utilizado para fins terminológicos em diferentes especialidades do conhecimento.

5.3 O léxico do português no Brasil e suas influências étnicas

A efetividade do léxico corresponde as formas como a língua foi e é moldada, e a resultante dos deslocamentos e dos encontros de seus falantes. A língua portuguesa falada no Brasil, nesse sentido, não se limita as origens dos portugueses, mas absorve a influência dos povos indígenas e dos povos africanos. A língua sob a forma do léxico, configura como um tecido do qual várias culturas convergem e inter cruzam a ponto de sustentar uma base linguística rica e diversificada, uma vez que cada parte desse tear se ajusta aos diferentes desenhos que representam suas particularidades transmitidas na fala e na escrita. É nesse ponto de encontro, que Biderman expõe a questão da formação e da consolidação do léxico na gênese linguística brasileira, pois o mesmo

[...] é a testemunha de uma cultura, porque é exatamente no domínio do léxico que as influências de substrato (indígena) e superestrato (africano) mais se manifestam [...]. Os conceitos evoluem acompanhando a marcha da sociedade e as suas mutações. E o vocabulário vai acompanhando e refletindo as transformações dos estados da sociedade. Assim é que, ao fazer a história das palavras, muita vez os linguistas fizeram simultaneamente a história da sociedade (Biderman, 1997, p. 308).

Os dois principais substratos mencionados por Biderman revelam os dois núcleos centrais dos aspectos étnico-raciais eclodidos no cenário de exploração colonialista presentes

na América Latina e no continente africano, mencionados por Quijano (1999; 2014) e Munanga (2004). O léxico nesse cenário, ao mesmo tempo em que está testemunhando a trajetória de uma sociedade, também é tido como o reflexo histórico e cultural.

5.3.1 Influências étnicas indígenas do Brasil

Segundo o último censo do IBGE (Instituto Brasileiro de Geografia e Estatística, 2022), atualmente, vivem cerca de 1.693.535 indígenas e cerca de 300 etnias. O território brasileiro tem origem dos antigos povos originários e descendentes das etnias resistentes que estiveram aqui antes da chegada dos portugueses, dentre outras nações e monarquias europeias. Estima-se que no século XVI, o território nacional abrigava mais de 2.000.000 indígenas (Instituto Brasileiro de Geografia e Estatística, 2000).

No início dos anos 2000, cerca de 181 línguas indígenas ainda eram faladas no Brasil, das quais pertencentes aos troncos linguísticos Tupi e Macro-jê – esse número é impreciso, considerando alguns casos de distinção entre línguas e dialetos¹⁵, além de duas ou três línguas que deixaram de ser faladas nos últimos cinco anos –, em contraste com os tempos passados de modo que se estimava a existência de 1,2 mil línguas faladas pelos povos indígenas (Rodrigues, 1986; 2005).

As línguas indígenas representam, além das origens linguísticas brasileiras, fatos sociais e formas de resistência das heranças culturais e étnicas. Os elementos expressos nessas línguas remontam traços do modo de vida e do imaginário cultural das famílias e do elo social, além da relação com o universo e a natureza, ambos próprios de culturas, filosofias, alquimias e conhecimentos apagados e ceifados pelas sociedades colonialistas e imperialistas.

Um dos princípios básicos que regem as línguas indígenas pode ser observado no entendimento do ser e sua relação direta com o universo, esse revelado nas diversas possibilidades de interpretação filosófica e no conjunto de histórias e memórias, das quais os etnólogos chamam de mitos. Esta talvez seja uma das barreiras encontradas para reconhecer os saberes étnicos diante da hegemonia dos colonos europeus, principalmente, os portugueses.

Como exemplo, Jecupé (1998) conta que a cultura Guanani tem o ato de nomeação tido como uma manifestação da parte céu de um ser para a sua parte terra, representando a tradição de quando não existia noite, chamada de “Arandu Arakuaa”, e a ciência chamada de

¹⁵ Segundo Saussure (1977) os dialetos são figurados como tipos linguísticos determinados e circunscritos, cobrindo territórios em justaposição. Em outros termos, o dialeto pode ser notado como uma variação da língua, assim em um âmbito nacional há possibilidade de existência de vários dialetos

“A Sabedoria dos Ciclos do Céu” ou “O Saber do movimento do Universo”. Esses fundamentos foram incorporados às palavras e tudo aquilo que seus indivíduos nomeiam, permitindo remetê-las aos espíritos, à criação e ao estabelecimento do universo.

Jecupé (1998) colabora ainda ao explicar que a maioria dos povos originários tiveram suas tradições do Sonho, do Sol e da Lua, ambas baseadas no Ciclo do Tupã – uma das quatro estações da natureza cósmica que fundamenta a ciência sagrada. O Ciclo do Tupã fornece aprendizados e justificativas da natureza e do entrelaçamento de todos os seres. As divindades são figuras representantes desse cenário de criação e entendimento do universo. A história antiga evidencia a gênese dos repertórios linguísticos e os saberes relativos a fala, como a compreensão dos Tubuguaçu (antigos ancestrais dos recentes povos Tupi-Guarani):

Cada vogal vibra uma nota do espírito que os ancestrais chamavam de angámirim, que comporta o *ayvu*, estruturando o corpo físico. São sete tons, e quatro deles referem-se aos elementos terra, água, fogo e ar, coordenando a parte física, emocional, sentimental e psíquica do ser. E três desses sons referem-se à parte espiritual do ser. Eis os tons: *ÿ* (espécie de "u" pronunciado guturalmente), *U* (vibrando da mesma maneira que o U da língua portuguesa), O, A, E, I (vibrando da mesma maneira que na língua portuguesa), e, por último, o som "insonoro", que não se pronuncia, mas que, na antiga língua abanhaenga, mãe da língua proto-tupi, se pronunciava unindo aproximadamente os sons mudos da expressão *MB*, gerando palavras como *Mbaekuaa*, *Mboray* (sabedoria, amor) (Jecupé, 1998, p. 24).

No trecho, o ser presente em cada tom possibilita a justificativa de alguns elementos na construção e na consolidação dos léxicos falados por esses ancestrais e seus descendentes. As vogais e as consoantes no contexto dessas línguas estreitam seus fundamentos nas histórias e memórias. Como exemplo, o Tupi-Guarani se destacou por ser um grupo genético e de moderada diferenciação interna da sua constituição, podendo ainda ser diferenciado por oito subgrupos fundamentados no pensamento desses povos (Rodrigues, 1986).

De acordo com características de manutenção e por serem vividas na oralidade, as línguas dos povos originários do Brasil tenderam a desaparecer com o avanço territorial desordenado e a ausência de condições sociais e culturais para a manutenção dos saberes. O desaparecimento dos povos que falavam as línguas anteriores à língua portuguesa foi ocasionado pelas consequências promovidas pelos europeus e seus descendentes através das campanhas de extermínio e caçadas pelos escravos e ou em detrimento das epidemias e doenças contagiosas (Rodrigues, 1986). Além do modo extrativista, exploração e cultivo, característicos da colonização nas Américas.

Uma das principais questões dos estudos linguísticos no Brasil está na interação entre duas línguas, nesse caso, a língua dos colonos e a língua dos colonizados. Braggio (1998)

cita que a interação entre duas línguas pode ser vista como uma transferência negativa, perspectiva behaviorista na Linguística, e um erro, do qual no ambiente escolar implicaria em rotulações, reprovação e repetição, embora na perspectiva dos racionalistas a língua não deve ser reduzida a imitação e sim adentrar no processo ativo e criativo. Isso não reduz os impactos da extinção de alguns grupos linguísticos e diferentes dos dialetos atuais.

Em análise histórica, Rodrigues (1986, 1993) relata que o maior número dessas línguas desapareceu nas áreas onde foram colonizadas de modo mais intenso e há mais tempo (área compreendida pelo Sudeste e pela maior parte das regiões Nordeste e Sul do Brasil), estimando que os vazios nas porcentagens de índices de falantes foram ocasionados pelo desaparecimento de um número considerável de línguas, provavelmente, pela ausência de algumas famílias inteiras. Assim atualmente, a presença indigenista linguística concentra-se em capitais da região norte, seguidas dos principais centros urbanos do Nordeste e do centro-oeste, diminuindo, consideravelmente, quando se aproximam daqueles localizados no Sudeste e Sul (Costa, 2012).

Outro importante fato a considerar é a interação as línguas e os aspectos étnicos dos grupos de africanos, ambos colocados em uma situação de exploração e escravização. Possivelmente, isso implicaria, além das relações distintas de contato social e até mesmo conflito, em modos e hábitos da comunicação. Por serem em maior número e estarem ambientados nos seus territórios de convívio, as etnias indígenas estiveram em maior número do que as africanas. Assim, as relações estabelecidas entre os portugueses e indígenas no Brasil, e consequentemente os processos e o estabelecimento da fala pelos seus sucessores e descendentes, geraram uma consequência linguística: o desenvolvimento das línguas gerais (Rodrigues, 1993).

Considerando um estudo acerca do léxico indígena nas capitais brasileiras diante da presença de outras culturas, como a espanhola e as etnias, dentre elas matrizes africanas, Costa (2012) relata o mapeamento de unidades lexicais referentes às atividades agropastoris, à fauna, à alimentação e à cozinha, demonstrando uma distribuição regional. A presença dessas unidades nessas cidades implica na proximidade de territórios remanescentes das tradições indigenistas e que, por sua vez, geraram um contato com a oralidade em português.

As palavras que nomeiam itens da flora e que constam em receitas para finalidades curativas não são os únicos exemplos da temática medicina indígena e saúde na atualidade. Gomes e Ferreira (2014) apontam a variedade de denominações da área da saúde mapeados em estudo etnoterminológico da medicina tradicional do povo Múndurukú, dentre os quais termos

do léxico que indicam doenças e possíveis tratamentos, remetendo aos nomes de plantas e às características sintomáticas.

Nesse sentido, as palavras utilizadas para denominar termos científicos da área da saúde são importantes indicativos de como essas línguas ainda deixam indícios. Entretanto, ainda é possível revelar a problemática notada na comunicação entre pacientes indígenas e profissionais da saúde, esses últimos detentores de protocolos e processos comunicativos estabelecidos exclusivamente em nível de especialidade e atualização (Bentes Pinto et al, 2020).

Por outro lado, os resquícios linguísticos, a partir das interações entre as línguas e as características da língua ou variedade adquirida, possibilitam a construção de uma variedade da língua portuguesa resultante da língua indígena e língua portuguesa. Conforme o caso da alfabetização do grupo Xerente e estudos da Sociolinguística intrinsecamente relacionado à educação escolar indígena (Braggio, 1998).

A distribuição dos itens lexicais dos indigenismos na área semântica das atividades agropastoris apresentam alguns exemplos encontrados nas capitais: como o manganá (ponta roxa do cacho da banana) é conhecido e falado somente nas regiões Norte e Nordeste; soca/soqueira (parte pequena que resta depois que se corta o pé de arroz ou de fumo) e a maniva e maniçoba (ambas do mesmo grupo semântico do cultivo e culinária da mandioca) conhecidas no Norte; o uru (objetos de vime, de taquara e de cipós trançados) e tamboeira (planta que apresenta grãos) mencionadas exclusivamente no Nordeste; e as variações de mandioca – aipim e macaxeira – amplamente mencionadas em todas as regiões brasileiras (Costa, 2012). Esses exemplos integram às temáticas de atividades cotidianas dos falantes e ampliam em referências dos elementos culturais que ainda permanecem vivos na oralidade e nas memórias documentadas.

5.3.2 Influências étnicas africanas: resistência e silenciamento da afro-brasilidade

Segundo o IBGE (2023), o Brasil recebeu entre os séculos XVI e XIX cerca de 4 milhões de homens, mulheres e crianças africanas, representando um terço de todo o comércio negreiro da época, o que simboliza a maior concentração de africanos deslocados de seu território de origem durante o período colonial. A trajetória histórica dos pilares políticos e nacionais brasileiros configurou-se na trajetória de exploração desses grupos.

Gonzalez e Ortiz (2022) recordam que essa trajetória foi marcada pelo sistema e os ditames do colonialismo e do imperialismo, sobretudo a partir do racismo como uma “ciência” que sustenta a superioridade euro-cristã (branca e patriarcal) e promove a violência e a

escravização dos africanos e seus descendentes. Esses aspectos marcaram não somente a chegada desses povos no território brasileiro, mas detiveram efeitos no desaparecimento e na manutenção de suas raízes culturais.

Uma das principais características desses povos se refere à diversidade cultural e étnica, demonstrando que o continente africano não era de todo modo um território heterogêneo. Isso é justificado pelo continente africano ser o berço da humanidade e desde as civilizações antigas, como as egípcias, até os impérios de Gana e Mali, desenvolvendo-se em diversos grupos e sob a riqueza de seus aspectos culturais (Munanga, 2009).

As etnias Ilorin, Ijexá, Abeokutá (egbás), Ketu e Ibadan, Lagos e Ifé originários das nações iorubás, chamados nagôs, são alguns dos remanescentes africanos no Brasil (Pradin, 2000). Esses povos descendem de comunidades situadas em determinadas regiões da África, correspondendo a determinadas famílias linguísticas. Ilustrando essa descendência, Cafundó e Tabatinga são duas comunidades negras que coincidem com o mesmo traço linguístico: o léxico de origem banto (Petter, 2005).

Além dos povos nagôs, Rodrigues (1935) recorda os últimos povos africanos no Brasil que deixaram influências culturais e religiosas de matriz afro: os Geges, os Minas, os Haussás, os Tapas (ou Mifês), os Bornus, os Gruncis (ou Galinhas); os Fulás (ou Filanins); os Mandingas (ou Mandês), os Bantus¹⁶ (dentre eles os Congos ou Cabindas, Angolas, Benguelas, Cassanges, Moçambiques, Macuas). Cada um desses povos apresentava traços étnico-culturais e características comportamentais diferentes, percebidas na fala, nas crenças religiosos, nas vestimentas e hábitos sociais.

A herança linguística africana foi marcada pela introdução de vocábulos e palavras para nomear e propagar as atividades sociais e as diversas manifestações artísticas, como a dança e a capoeira, e as atividades de cultivo de plantações e culinária. Munanga (2009) revela algumas manifestações e seus nomes a partir dos elementos culturais dos bantos presentes na dança e na música: congos, quilombos, coco, jongo, maculelê, maracatu, bumba-meu-boi e capoeira. Além da repercussão e da fundação de gêneros tipicamente afro-brasileiros como o samba. Com o reflexo da escravização e da violência, esses grupos estiveram submetidos ao processo de integração com outros indivíduos: os falantes da língua portuguesa e das línguas indígenas.

A segunda reflexão diz respeito às línguas africanas que foram trazidas da África para o Brasil. Qualquer que seja sua identidade e sua antiguidade, trata-se sempre de

¹⁶ ou bantos

línguas desenraizadas de seu nicho ecológico, submetidas, além disso, ao perigo de diversas rupturas. Daí decorre seu estado de degradação. A ruptura semântica foi, sem dúvida, a que mais se sentiu: para os africanos deslocados, o sentido das palavras tornou-se brutalmente obsoleto ou passou a "girar em falso", porque elas não refletiam mais a realidade africana, mas, ao mesmo tempo, ainda não tinham apoio na nova realidade, constituída de noções diferentes e de denominações novas (plantas, farmacopeia, caça, animais, novas técnicas e novos produtos de consumo). A outra ruptura foi de ordem dialógica, pois essas mesmas línguas foram confrontadas com contatos linguísticos inabituais: novas línguas convivendo no mesmo espaço (outras línguas africanas decorrentes do tráfico, línguas ameríndias, língua portuguesa falada pelos senhores), sem esquecer a perda de seu valor identitário consecutivo ao nivelamento da língua resultante do desaparecimento das variantes dialetais (Bonvini, 2014, p. 33).

Nesse trecho de Bonvini, são compreensíveis o conflito e as consequências negativas da ruptura daqueles que tinham as línguas africanas como sua principal qualidade e ligação com suas origens. O que, por outra perspectiva, impactou na participação histórica e nas características da língua portuguesa falada atualmente no território brasileiro.

Mendonça (2012) afirma a participação efetiva dos negros africanos na formação fonética do português falado no Brasil, pois esses aprenderam a falar o português traduzindo suas ideias ao partir do modo de falar africano. Apesar dessa influência ser de certo modo concorrente da herança das línguas indígenas, a participação na fonética e na morfologia de algumas palavras se repercute até o presente em diversas regiões. Diante disso, Rougé (2014, p. 68) recorda que “do ponto de vista lexical, chama atenção o fato de que os termos de origem africana são pouco numerosos e geralmente pertencem a domínios bem específicos: vegetação, culinária, "cultura"”. Domínios esses deixados como uma das poucas opções das quais esses descendentes brasileiros poderiam expressar suas origens.

Dentre as línguas, o Tabatinga destacou-se como uma língua adquirida na juventude na faixa etária entre 11 e 20 anos, principalmente entre amigos e não foi transmitida como língua materna, e o Cafundó, tida como uma variedade do português regional (um dialeto rural), caracterizada por um léxico reduzido de origem banto (quimbundo, em particular) com o foco no uso lúdico e com a finalidade de ocultação (Petter, 2005). As etnias africanas, nesse sentido, implicaram em padrões linguísticos de adaptação à medida em que os processos de colonização e catequização dos Jesuítas se efetivaram nos processos de dominação.

Esses grupos advindos de diferentes culturas e distintas regiões africanas estabeleceram-se conforme o cenário de luta e sobrevivência a que foram submetidos nas faixas de exploração e colonização americana. Nesse cenário, as heranças africanas dos quilombos permaneceram no Brasil, representando traços da existência das comunidades rurais predominantemente negras (Petter, 2005; Munanga, 2009).

Alkmim e Petter (2014) apresentam um levantamento de vocábulos de origem africana (Quadro – 6) conferidos nos registros de estudiosos dos séculos XIX (1826-1897) e XX, do qual a exposição revela palavras amplamente difundidas no léxico brasileiro conforme a amostragem do corpus coletado a partir do estudo “Falares africanos na Bahia” (um vocabulário afro-brasileiro) de Yeda Pessoa de Castro:

Quadro 6 – Registro de vocábulos de origem africana no português brasileiro: século XX

Categorias	Termos
Categoria de termos utilizados em interações sociais	banzo- 'tristeza', 'saudades de pessoas e lugares familiares'; caçamba - 'depósito de lixo' e 'carroceria de caminhão'; cachaça - 'aguardente', 'pinga'; cachimbo - 'aparelho para fumar'; caçula- 'filho mais novo'; candango - 'trabalhador que construiu Brasília'; canga - 'instrumento para prender o pescoço dos animais'; capanga - 'bolsa' e 'guarda-costas'; carimbo- 'marca e instrumento para marcar'; caxumba - 'parotidite'; cochilar- 'dormitar'; corcunda- 'corcova, giba'; denço - 'manha, comportamento infantil'
Categoria de termos informais (de uso coloquial que, em algumas situações são substituídos)	bamba - 'exímio', 'hábil'; bambambã- 'o mais enrendido', 'o especialista'; banguela - 'desdentado', 'falta de dente'; cafuné - 'carinho feito na cabeça com a ponta dos dedos'; carimba/catimbeiro- 'manha', 'astúcia'; catinga - 'mau cheiro'; mandinga - 'feitiço'; muamba - 'contrabando'; muxoxo- 'gesto de desdém com estalo de língua no céu da boca'.
Categoria de termos informais (de uso restrito)	angu - 'confusão' 10; babaca- 'bobo'; babau- 'acabou!'; biboca- 'habitação ou local ruim'; bunda - 'nádegas'; cafofo - 'esconderijo'; cafundó- 'lugar muito distante'; cambada- 'bando de vagabundos, corja'; cucuia - 'malogro, acabou!'; muquifo- 'habitação muito ruim, precária, suja'; 'biboca'; muquirana- 'avarento', 'pão-duro'; muvuca- 'confusão, algazarra, mistura de coisas ou de pessoas'; muxiba - 'pelanca'; guizumba- 'confusão'; sacana - 'mau caráter'; ziquizira - 'azar'; 'doença indefinida'; zumbi- 'alma penada'; 'maldormido'

Fonte: Adaptado de Alkmim e Petter (2014, p. 157-158).

Em outros exemplos, há o caso dos adjetivos no dialeto local dos estados de Pernambuco e Alagoas: *capiongo*, *cassange*, *cafuçu*, *ingangento*, *cangulo*, *macambúzio*, *manzanza*, *caçula*, *buzuntão*, *capenga*, *banguelo*, *fiota*, *dunga*, *zorô*, *granganzá* e *cutuba* (Mendonça, 2012). Diante desses elementos herdados, os empréstimos lexicais são conhecidos como africanismos e seus aspectos se expandem para traços fonológicos, morfológicos e sintáticos (Bonvini, 2014).

Corroborando, Mendonça (2012) analisou e apontou algumas alterações fonéticas de origem africana a partir dos respectivos exemplos: vocalização - a mudança do fonema *lh* para a semivogal *y* (mulher = *muyê*, melhor = *meyor*); assimilação - fonema *j* para o sibilante *z* (José = *Zezé*, registro = *rezisto*); dissimilação - a ocorrência nos grupos consonânticos de elocução difícil (negro = *nego*, alegre = *alegue*); aférese - presente em *tá* = *estar*, *ocê* = *você*; *cabá* = *acabar*; *apócope* - presente em *general* = *generá*; *cafezal* = *cafezá*, *mel* = *mé*.

Outras classes gramaticais também receberam sua herança, como alguns verbos de primeira conjugação: *mandingar*, *zangar*, *bongar*, *carimbar*, *catingar*, *banzar*, *sambar*, *curiar*, *maxixar*, *cochilar*, *candongar*, *enquizilar*, *aquilombar* (Mendonça, 2012).

Atualmente, as línguas africanas manifestam-se como línguas especiais e no seu uso de língua secreta, em função da manifestação dos chamados cultos afro-brasileiros e ou em função de demarcação social (Petter, 2005). Nessa perspectiva, é importante reconhecer que a efetivação aconteceu e acontece não somente por meio da oralidade e nos poucos registros que se tem, mas adquire sua permanência na resistência dos hábitos culturais e na essência que o ato da fala denota a sobrevivência. Isso acontece considerando que “[...] as relações construídas pelos contatos culturais se efetivam e ganham referenciais por mediação do léxico, sobre o qual incidem os percalços dos processos interculturais e dos atos tradutórios” (Caldas; Barros; Silva, 2016, p. 121).

A partir desse cenário, entende-se que a permanência e o desenvolvimento das línguas africanas encontraram no racismo e na violência seus primeiros instrumentos de destruição e de silenciamento. O falar dos negros africanos até os dias atuais é tido como um tipo de código secreto “[...] como meio de ocultar as conversas, principalmente em presença de estranhos. É provável que essas línguas tenham exercido essa mesma função no passado a fim de esconder dos senhores as palavras trocadas entre os escravos, sobretudo quando eles planejavam fugas. Hoje elas continuam a exercer esse papel, mas, frequentemente, transpostas para outros conflitos sociais” (Bonvini, 2014, p. 51).

Os tipos de racismo instalados (racismo manifesto, racismo por negação) na política e na própria constituição da sociedade brasileira estiveram pautados em diferentes níveis e

diferentes formas de opressão racista, dentre as quais as hegemonias no poder político e na ciência através das epistemologias (Gonzalez; Ortiz, 2022). Essa recorrência histórica, impediu e impede de certo modo na transmissão e na construção linguística e cultural diante do silenciamento e da descaracterização das etnias africanas desde o período colonial.

6 ANÁLISE E DISCUSSÃO

De acordo com os objetivos específicos e em função de responder ao objetivo geral desta pesquisa, as seções deste capítulo analisam e discutem as categorias, de modo, que os trechos e citações relevantes foram cruciais para estabelecer diálogos entre os autores e, conseqüentemente, pautar as inferências aqui relatadas.

6.1 Aspectos conceituais

O Quadro 7 a seguir menciona os principais conceitos mapeados durante as buscas combinadas nas temáticas contempladas:

Quadro 7 – Aspectos conceituais

Conceitos	Análise de Domínio	Tesauros	Léxico	Etnia, Raça, Cultura e Análise de Domínio	Tesauros e Análise de Domínio	Etnia, Raça, Cultura e Tesauros
BIAS	Feinberg, Melanie - 2007	Milani, Suellen Oliveira; Guimaraes, Jose Augusto Chaves – 2010				
Classificação social	Albuquerque, Ana Cristina – 2023			Albuquerque, Ana Cristina – 2023		
Comunidade discursiva	Hjørland, Birger; Albrechtsen, Hanne – 1995; Nascimento, Denise M.; Marteleto, Regina Maria – 2008; Hjørland, Birger – 2017; Amorim, Igor Soares; Café, Lúcia – 2017; Wang, Lin; Qiu, Junping – 2021	Mai, Jens- Erik – 2008		Nascimento, Denise M.; Marteleto, Regina Maria – 2008	Mai, Jens- Erik - 2008	
Competência lexical			Guerrero, Gladys Yolanda Pasuy; Montoya, Claudia Liliana Agudelo - 2022			

Conceitos	Análise de Domínio	Tesaurus	Léxico	Etnia, Raça, Cultura e Análise de Domínio	Tesaurus e Análise de Domínio	Etnia, Raça, Cultura e Tesaurus
Contexto	Mai, Jens-Erik – 2005	Mai, Jens- Erik – 2008			Mai, Jens-Erik - 2008	
Cultura	Rodrigues, Anderson Luiz Cardoso - 2012; 2015 Barité, Mário; Rauch, Mirtha - 2020			Rodrigues, Anderson Luiz Cardoso - 2012; 2015 Wang, Shengang – 2019; Barite, Mário; Rauch, Mirtha – 2020		
Disciplina	Mai, Jens-Erik - 2005	Mai, Jens- Erik – 2005				
Domínio	Hjørland, Birger; Albrechtsen, Hanne - 1995; Mai, Jens-Erik - 2005; Albrechtsen, Hanne - 2015 Tennis, Joseph T. - 2012 Hjørland, Birger - 2017; Damus, Maria Arminda; Acuna, Gabriela Noemi - 2019;	Hjørland, Birger – 2002a; Mai, Jens- Erik – 2008			Birger Hjørland - 2002; Mai, Jens-Erik - 2008	
Etnoconhecimento	Moraes, Miriam Gontijo de - 2018	Moraes, Miriam Gontijo de -2018		Moraes, Miriam Gontijo de - 2018		
Garantia cultural	Alvarenga, Lídia; Dias, Célia da Consolação – 2012; Barite, Mário; Rauch, Mirtha - 2020	Moura, Maria Aparecida; Assis, Juliana - 2010 Gomes, Pablo; Frota, Maria Guiomar da Cunha – 2019; Coladangelo, L. P. - 2021		Coladangelo, L. P. - 2021		Oliveira, Dalgiza Andrade; Araújo, Ronaldo Ferreira de - 2012
Garantia de uso		Moura, Maria Aparecida – 2009; Moura, Maria Aparecida; Assis, Juliana - 2010				Oliveira, Dalgiza Andrade; Araújo, Ronaldo Ferreira de - 2012
Garantia estrutural		Moura, Maria Aparecida - 2009				Oliveira, Dalgiza Andrade; Araújo, Ronaldo Ferreira de - 2012

Conceitos	Análise de Domínio	Tesaurus	Léxico	Etnia, Raça, Cultura e Análise de Domínio	Tesaurus e Análise de Domínio	Etnia, Raça, Cultura e Tesaurus
Garantia filosófica		Gomes, Pablo; Frota, Maria Guiomar da Cunha - 2019				
Garantia literária	Beghtol, Clare – 1995; Barite, Mario; Fernandez-Molina, Juan Carlos; Guimaraes, Jose Augusto Chaves; Ernesto de Moraes, Joao Batista – 2010; Alvarenga, Lídia; Dias, Célia da Consolação – 2012; Coladangelo, L. P. - 2020	Moura, Maria Aparecida – 2009; Moura, Maria Aparecida; Assis, Juliana – 2010; Gomes, Pablo; Frota, Maria Guiomar da Cunha - 2019		Coladangelo, L. P. - 2020		Oliveira, Dalgiza Andrade; Araújo, Ronaldo Ferreira de - 2012
Garantia semântica		Gomes, Pablo; Frota, Maria Guiomar da Cunha - 2019				
Gênero documental	Hernandez, Bianca Ferreira; Vital, Luciane Paula - 2022					
Questões de gênero	Pinho, Fábio Assis; Guimarães, José Augusto Chaves - 2012; Albuquerque, Ana Cristina - 2023	Olson, Hope A. - 2007; Moraes, Miriam Gontijo de - 2014; Ribeiro, Ana Rosa Pais; Decourt, Beatriz; Almeida, Tatiana - 2017; Sousa, Brisa Pozzi de; Tolentino, Vinicius de Souza – 2017	López-Huertas, María J. - 2007	Albuquerque, Ana Cristina – 2023		
Hospitalidade cultural	Coladangelo, L. P. - 2021; Barite, Mario; Rauch, Mirtha - 2020			Coladangelo, L. P. - 2021		Coladangelo, L. P. – 2021
Interseccionalidade	Coladangelo, L. P. - 2021			Coladangelo, L. P. - 2021		Coladangelo, L. P. – 2021

Conceitos	Análise de Domínio	Tesauros	Léxico	Etnia, Raça, Cultura e Análise de Domínio	Tesauros e Análise de Domínio	Etnia, Raça, Cultura e Tesauros
Léxico		Engerer, Wolkmar – 2017	Šnajder, Jan; Bašić, B. Dalbelo; Tadić, Marko - 2008; Bentes Pinto, Virginia; Campos, Henry de Holanda; Aragão, Maria do Socorro; Ribeiro, Cristina da Costa – 2020			
LGBTQI (Lésbica, Gay, Bissexual, Transgênero, Queer e Intersexo)		Golub, Koraljka; Bergenmar, Jenny; Humelsjo, Siska – 2022				
LGBTQIAP+ (Lésbicas, Gays, Bissexuais, Transexuais/Trans gêneros/Travestis, Queer, Intersexual, Assexual, Pansexual e as demais orientações e identidades)		Bissoli, Bruna da Silva; Covello, Lucas Gatto; Pisseli, Bianca Iris; Santos, Raphael Augusto – 2018				
Colonialidade do poder	Albuquerque, Ana Cristina – 2023			Albuquerque, Ana Cristina – 2023		
Raça	Albuquerque, Ana Cristina – 2023	Ribeiro, Ana Rosa Pais; Decourt, Beatriz; Almeida, Tatiana – 2017		Albuquerque, Ana Cristina – 2023		

Conceitos	Análise de Domínio	Tesauros	Léxico	Etnia, Raça, Cultura e Análise de Domínio	Tesauros e Análise de Domínio	Etnia, Raça, Cultura e Tesauros
Relações semânticas		Mazzocchi, Fulvio; Tiberi, Melissa; De Santis, Barbara; Plini, Paolo – 2007; Stock, Wolfgang G. – 2010; Hjørland, Birger – 2013b, 2015; Bräscher, Marisa – 2014; Maculan, Benildes Coura; Lima, Gercina Ângela Borém de Oliveira – 2014; Weiss, Leila Cristina; Bräscher, Marisa – 2016; Souza, Jacqueline A.; Pestana, Olivia – 2017; Almeida, Maurício Barcellos; Emygdio, Jeanne Louize – 2019; Nunes, Flávia Rodrigues Elias; Maculan, Benildes Coura Moreira dos Santos; Almeida, Maurício Barcellos – 2021				
Vocabulário		Moreiro, Jose Antonio – 2018	Bentes Pinto, Virginia; Campos, Henry de Holanda; Aragão, Maria do Socorro; Ribeiro, Cristina da Costa – 2020			
Negritude		Santos, Andréia dos; Costa, Amabile; Barros, Camila Monteiro; Vital, Luciane Paula – 2020				Santos, Andréia dos; Costa, Amabile; Barros, Camila Monteiro; Vital, Luciane Paula – 2020
Afro-brasileiro		Santos, Andréia dos; Costa, Amabile; Barros, Camila Monteiro; Vital, Luciane Paula – 2020				Santos, Andréia dos; Costa, Amabile; Barros, Camila Monteiro; Vital, Luciane Paula – 2020

Fonte: Elaborado pela autora.

O termo **domínio** é amplamente mencionado e associado à delimitação de objeto de pesquisa, ao recorte teórico e à caracterização de uma temática referentes aos sujeitos de um estudo. Mai (2005) aponta a ausência de explicações detalhadas do conceito de “domínio” em Hjørland e Albrechtsen (1995). Nesse sentido, Mai realoca o conceito na discussão das abordagens e na prática de indexação, tendo em vista que o trabalho do indexador pode ser guiado conforme a importância ou a articulação dada por uma abordagem.

Contudo, Mai (2005) ao apresentar o domínio como um conceito aberto e em evolução, indicou que esse é utilizado para se tratar de um grupo de pessoas que compartilham objetivos comuns, uma área de especialização, um conjunto da literatura ou um grupo de colaboradores de uma organização. Essa noção denota um aspecto conceitual relevante: aquele que reconhece a integração entre pessoas e documentos no seu ambiente de prática informacional. O que implica em uma proposta destinada a superar a fixação na abordagem do assunto, ou seja, tratar somente o documento e, no máximo, utilizar o respaldo de especialistas para efetivar e avaliar os processos e os instrumentos utilizados na representação temática. Mai, aqui, ainda se aproximou do conceito de Smiraglia (2012) do qual um domínio é um grupo com uma base ontológica com hipóteses comuns, consenso epistemológico e semântica social, antes mencionado no referencial teórico desta pesquisa.

Tennis (2012) revelou que o domínio deve ser declarado em termos de definição, escopo, alcance e finalidade. Entretanto, não é direcionado a ideia central do que é de fato um domínio, mas o autor aponta esta declaração que parece ser uma fórmula para definir qualquer domínio tratado na Análise de domínio, na qual ele exemplifica com o caso do domínio Shaker-ísmo.

Percebe-se nessa ideia de Tennis que a análise deve iniciar com a designação da identificação, do escopo, do alcance e da finalidade, o que condiciona o analista a se dedicar à visão determinista e antecipada do domínio que deverá ser descoberto em termos ontológicos. Lembrando que a dimensão ontológica, apontada por Hjørland e Hartel em 2003, se destaca como um caminho bem mais promissor para definir um determinado domínio. Quanto à identificação, escopo e finalidade, esses podem ser parâmetros norteadores da metodologia aplicada, cujos detalhes deverão servir de prospecção em relação aos limites e à delimitação da análise.

Em uma interpretação diferente daquela de Tennis em 2012, Mai (2008) considerou esse conceito como algo em desenvolvimento e aberto, na medida em que se desenvolverá o conceito do domínio, do qual ele é usado e aplicado na pesquisa e na prática. Observa-se ainda

uma reflexão de que o conceito do termo “domínio” é elementar para abstrair o conceito de um determinado domínio, como o da cultura organizacional, e o do “fato social” alocado nas teorias sociológicas europeias.

Albrechtsen em 2015, ao apresentar uma alternativa metodológica diante do paradigma do processamento da informação, tentou visualizar a concepção de domínio no desenvolvimento de software. A autora, ainda, salientou que não existe um domínio “pronto a mão” já que “os domínios não são terrenos lá fora, à espera de serem descritos e analisado pelos poucos iniciados. Fundamentalmente, todos nós podemos criá-los” (Albrechtsen, 2015, p. 561, tradução nossa). Este estudo não se firma como os demais estudos da AD na literatura em Ciência da Informação. E a implicação dada aqui não desenha um conceito claro. Porém, Albrechtsen tem como objeto os elementos a serem reutilizados na elaboração de um determinado software. Nesse caso, pode-se indagar, conforme a etimologia de “domínio”¹⁷, se tal objeto é a soma de um conjunto de variáveis ou algo a ser dominado perante regras e propósitos estabelecidos.

Hjørland (2002a; 2017) assinalou apontamentos esclarecedores baseados em preposições epistemológicas e resultado de uma revisão teórica e diálogos de outros autores, chegando, assim, a um denominador comum. Cabe destacar as contribuições de Hjørland como influentes na CI e na Biblioteconomia, o que desencadeou uma receptividade e um consentimento relevante da maioria dos membros da comunidade de pesquisadores, apesar de uma parcela de pesquisadores apresentarem contrapropostas e discordarem de alguns pontos, como o caso do não desenvolvimento da noção de domínio nos primeiros artigos.

Desse modo, a noção de domínio e observações são respondidas no artigo de 2017 “Domain Analysis” presente na seção da *Knowledge Organization* e enquanto verbete da Encyclopedia of Knowledge Organization da ISKO. Nos trechos abaixo o autor argumenta e disserta de modo que aponta para indagações epistemológicas essenciais para reconhecer a ideia de domínio:

O domínio é o conjunto total de informações para o qual, idealmente, espera-se que haja uma resposta para um problema. Em particular, se o problema é aquele que exige uma "teoria" como resposta, o domínio constitui o conjunto total de informações que

¹⁷ Hjørland (2017) recorda os cinco sentidos do substantivo “domínio” de acordo com o banco de dados WordNet: 1. Esfera, domínio, área, órbita, campo, arena; 2. Domínio, terra (território, o qual se exerce a regra ou o controle); 3. Domínio (o conjunto de valores da variável independente); 4. Mundo, domínio (pessoas em geral, um grupo distintivo de pessoas que compartilham interesses); 5. Domínio, região, domínio (um domínio do conhecimento).

devem, idealmente, ser explicadas por uma teoria da qual resolve esse problema” (Shapere, 1984 *apud* Hjørland, 2017, p. 439, tradução nossa).

Um domínio é um corpo de conhecimento, definido social e teoricamente como o conhecimento de um grupo de pessoas que compartilham compromissos ontológicos e epistemológicos. Os domínios são muitas vezes disciplinas acadêmicas, mas também podem ser, por exemplo, *hobbys*. (Hjørland, 2017, p. 6, tradução nossa).

É possível notar a visão do conceito de Shapere mencionado por Hjørland está a ideia de domínio que recorre ao conceito de teoria. Além disso, nota-se que as informações são tomadas como representação das análises de domínio e do processo de articulação e comprovação de uma teoria, no tocante daquilo que alimenta e daquilo que é produzido.

Outra questão à luz das observações de Hjørland: o domínio é um construto ontológico e epistemológico que pode sofrer atualizações e pode ser qualificado de acordo com sua evolução, sendo este ou não uma disciplina. Nesse sentido, há aproximação com o conceito de **disciplina** expressado por Mai (2005) quando apontou as seguintes colocações:

“Uma disciplina é definida como “um conjunto de especialidades” (Dogan, 2001, p. 14851) que representa “agregados evolutivos históricos de interesses acadêmicos compartilhados” (Chubin, Porter, & Rossini, 1986, p. 4) e as atividades são organizadas em disciplinas principalmente porque compartilham pontos históricos comuns. Os limites das disciplinas não são claros no sentido de que o que para uma pessoa pode ser sociologia será para outra política ou história; as pessoas estão colaborando entre disciplinas, publicando fora de suas próprias disciplinas e extraindo conhecimento de diversas disciplinas (Mai, 2005, p. 606, tradução nossa, grifo nosso)

Nesse raciocínio, a suposição de Mai está na especialização do conhecimento e no caráter colaborativo durante seus processos. Essa qualidade de especialização é notada na teoria dos campos e no estabelecimento das ciências na visão de Bourdieu (2004a; 2004b¹⁸) ao explicar uma definição de disciplina, na concepção de campo como um *nomos* particular e princípio de visão e divisão. É nesse, ainda, que a especificidade do campo acontece com os princípios de verificação de conformidade do real. O conceito de disciplina nesse parecer é exclusivo para o âmbito científico, entretanto, quando abordado no contexto não científico, suas características e suas estruturas são diversificadas e por vezes desconhecidas.

Com Hjørland, a possibilidade de ser um *hobbie* (*hobby*) pode resgatar outra perspectiva com viés em outros tipos de conhecimentos, assim podemos questionar: um

¹⁸ Cada disciplina (como campo) é definida por um *nomos* particular, um princípio de visão e divisão, um princípio de construção da realidade objetiva irreduzível ao de outra disciplina - segundo a fórmula de Saussure: «o ponto de vista cria o objeto» (o carácter arbitrário deste princípio de constituição que é constitutivo do «olho disciplinar» resume-se no facto de se enunciar, em geral, na forma de tautologias, como por exemplo, na sociologia, «explicar o social pelo social», ou seja, explicar sociologicamente as coisas sociais) (Bourdieu, 2004b, p. 75).

conhecimento popular ou qualquer outro interesse não acadêmico e científico é candidato a receber tal nomeação de “domínio”?

Se formos direcionar essa pergunta ao produto ou representação de um grupo étnico brasileiro, seja registrado em documento ou presentes na oralidade, como o caso de receitas medicinais ou objetos artísticos, o carácter ontológico reafirmado pelos membros e a epistemologia embasada nos saberes ancestrais podem, assim, serem equiparados ao *modus operandi* de uma disciplina científica. Esse paralelo enxerga o caso das temáticas analíticas de domínio que envolvem outras epistemologias que não aquelas elaboradas em tradições científicas europeias, mas nota a proposta de um domínio ser um objeto de estudo de outras epistemologias: dentre eles conceitos de gênero reformulados ou elaborados em vertentes como o feminismo negro latino e o transfeminismo latino.

O conceito de **contexto** foi ressaltado de modo a comprovar significância, tanto na temática de domínio quanto na questão de construção de tesouros, sendo imprescindível para formulação de qualquer sistema de classificação e organização não universal ou especializado. Para Mai (2005, p. 605, tradução nossa) “o contexto é frequentemente definido como aquilo que envolve uma palavra, passagem, evento ou situação específica. Para orientar os indexadores no processo de indexação, é necessário operar com uma noção mais precisa de contexto”. É comum o conceito ser associado ao espaço originário de um discurso, daí se localiza o texto, ou seja, o documento e sua procedência histórica, discursiva e temática. Nesse sentido, os gêneros documentais podem ser agregadores desse contexto remetido à indexação.

Hernandez e Vital em 2022 explicou o **gênero documental** a partir dos estudos e da Teoria de Gênero, justificando que os estudos de gêneros foram uma das 11 abordagens reconhecidas por Hjørland para o tratamento de acervos documentais responsáveis por subsidiar a metodologia destinada à descoberta de domínios e à produção informacional de suas comunidades discursivas. Esse conceito esteve fixado no cenário dos paradigmas e abordagens pós-custodial e pós-moderna ambientados na Arquivologia.

Essas autoras reconheceram que os estudos de gêneros, além de convergir diálogos de algumas disciplinas, tiveram influência e origens na Teoria de Gênero e influenciada por outras Teorias e Abordagens, como a Teoria da Estruturação de Giddens; os estudos de gêneros do discurso apontados por Mikhail Bakhtin; e a Teoria da Atividade de Bazerman, Russel e Winsor (Hernandez; Vital, 2022). Analisando, o gênero documental não é reduzido ao significado de tipologia e especificação textual, mas é compreendido enquanto forma de representação dos discursos situados nas atividades sociais e comunicativas.

É necessário destacar que os estudos de gêneros documentais mencionados pelos autores analisados não citam gêneros documentais e literários focados em textos ou comunicações orais étnicas ou advindas de contextos não científicos.

O artigo “*Adaptación de los Vocabularios documentales Al ambiente digital en red: léxico, significado y relaciones semânticas*” de 2018, Moreiro aponta características que embala, conceitualmente, a ideia de **vocabulário** e outras distinções concebidas de acordo com funções e usos, são eles: os vocabulários terminológicos, os vocabulários conceituais completos e os vocabulários como redes semânticas. A partir desse estudo, pode-se considerar os vocabulários como um conjunto de unidades que expressam as línguas e as linguagens em seus contextos. E com Moreiro, interpreta-se a existência de vocabulários de acordo com seu uso, por exemplo, no caso dos terminológicos, é tido como parte integrante da comunicação em especialidades.

Diferente do conceito de vocabulário, o **léxico** segundo Pinto, Campos, Aragão e Ribeiro (2020, p. 183) é entendido como “[...] o conjunto de todas as palavras de uma língua e, portanto, em manifestações concretas utilizadas nos processos de comunicação dos indivíduos de uma comunidade ou grupo social”. Os autores defenderam seu posicionamento no contexto de utilização dos léxicos na construção de vocabulários controlados.

Šnajder, Bašić e Tadić em 2008, não definem léxico, embora o introduziram no tema da utilização de uma abordagem de lematização e *stemming* – ambas abordagens fixadas na vetorização para seleção e otimização de palavras – na aquisição automática de léxico flexional para normalização morfológica.

Engerer (2017) comparou os léxicos aos tesouros e relatou que ambos são coleções estruturadas de signos, em particular palavras, embora diferem crucialmente em relação às propriedades comunicativas e semióticas. Os léxicos são constituídos nesse sentido de unidades da linguagem e das línguas naturais, já os tesouros as unidades da linguagem especializada. Porém, todas essas unidades derivam de sistemas formalizados por convenções, dentre as quais atentam para as características morfológicas, semânticas e outras.

Outro conceito presente nas discussões de Guerrero e Montoya, em 2022, acerca dos estudos léxicos na América Latina é a **competência lexical** sob o olhar da teoria linguística semântica e os estudos da Sociolinguística americana, hispano-americana e colombiana. Nesse sentido, compreende-se que a competência lexical é a capacidade de associação dos significados e as demais associações nos processos comunicativos e de retificação do léxico, dentre as quais conceituais e semânticas.

Nas investigações dos tesouros, as **relações semânticas** foram avaliadas enquanto imprescindíveis para o entendimento das relações pragmáticas e das sintagmáticas. Os autores abaixo exploram conceitualmente:

Relações semânticas são associações significativas entre entidades, de interesse para a Ciência da Informação, pelo papel essencial que desempenham nos processos que pessoas usam para representar conhecimento. Dentre as relações semânticas, destaca-se uma relação genericamente denominada por parte-todo, que lida, como seu nome sugere, com entidades tratadas como todos e as partes ligadas a elas (Almeida; Emygdio, 2019, p. 32).

As **relações semânticas são estabelecidas através da análise das características ou propriedades dos conceitos**, que nos permitem identificar diferenças e semelhanças que evidenciam determinados tipos de relações[...]. Dois conceitos conectados por meio de um relacionamento podem ser representados como um tríptico: [conceito] → (relação) → [conceito] (Bräscher, 2014, p. 176, tradução nossa).

Essas definições situadas na Organização do Conhecimento podem ser revisitadas a partir das influências das Teorias Semânticas e da Linguística Estrutural. A escola tradicional da Linguística Estrutural de Ferdinand de Saussure fixada na estrutura de um sistema onde podemos compreender os termos em relações sintagmáticas (no plano horizontal pertencente aos sintagmas) e paradigmáticas (no plano vertical da semântica) (Saussure, 1977).

Outro ponto central está na ideia de que as relações são conexões aprimoradas na condição de significação e de entendimento das unidades lexicais e suas identidades previstas em cada sistema linguístico e terminológico. Nesse interim, Stock em 2010, Almeida e Emygdio em 2019, Bräscher em 2014 e Weiss e Bräscher em 2016 retrataram esse conceito e as funções dessas relações no desenvolvimento dos sistemas de organização do conhecimento com base em relatos teórico-metodológico. Ambos demonstraram estudos mais críticos diferente das aplicações técnica e prática descritas nas normas ISO 25964-1 de 2011 e a ISO 25964-2.

Hjørland (2013b) afirmou que as relações semânticas são relações de significado, ou seja, relações estabelecidas por conceitos, e ambas são típicas de tesouros e sistemas de classificação. Por outro lado, Mazzocchi, Tiberi, Santis e Plini (2007) enfatizaram as teorias semânticas e a consistência das relações semânticas destinadas a orientar usuários nos caminhos da recuperação da informação, destacando como possibilidade: introduzir no quadro de construção de tesouros as noções de jogos de Linguagem de Ludwig Wittgenstein e do horizonte hermenêutico.

Há dois tipos de relacionamentos semânticos: as relações paradigmáticas – dentre as quais hierárquicas, equivalentes e associativas ou específicas – e as relações sintagmáticas (Hjørland, 2015; Stock, 2010). Hjørland (2015) mencionou que as relações paradigmáticas

também são denominadas de “associativas” e tem seu conceito originário nas investigações do semiótico e linguista Saussure; elas apoiam-se em ligações traçadas por padrões mentais e culturais e agrupamentos linguísticos por afinidade ou analogia; além de dependerem do contexto e de definições e serem intimamente ligadas aos paradigmas e teorias de cada disciplina e ramo do conhecimento. O trecho abaixo exemplifica de modo prático:

Na frase “JASIST é uma revista em Ciência da Informação” todas as palavras têm uma relação sintagmática. Eu poderia ter dito, no entanto, “JASIST é uma revista em ciência da computação” (ou “em Bibliometria” ou “em Gestão de Informação”, etc.). As relações entre Ciência da Informação, Ciência da Computação, Bibliometria e Gestão da Informação são, nesta Terminologia, entendidas como relações paradigmáticas (Hjørland, 2015, p. 1367, tradução nossa).

Interpretando essas relações semânticas, pode-se considerar que as paradigmáticas são relacionamentos estabelecidos entre unidades linguísticas e terminológicas de significado, conforme a substituição ou associação entre outras unidades de mesma categoria (substantivos, verbos, sintagmas ou locuções). As sintagmáticas se referem aos relacionamentos encadeados no arranjo das unidades presente em uma frase ou sintagmas compostos, do qual a composição e o posicionamento de cada unidade podem ser invertidos ou combinados.

Ambas devem ser compreendidas a partir das normas e das estruturas de cada sistema linguístico. Além disso, as normas e os elementos interpretativos solicitam uma investigação das heranças relatadas pelas escolas Linguísticas, das quais estão presentes na elaboração de vocabulários e linguagens, como os tesouros e as listas de cabeçalhos de assuntos. Contudo, essas relações detêm características semelhantes ou são estabelecidas em sistemas como aqueles das línguas indígenas da América Latina, como o Tupi-guarani, ou de línguas africanas, como o Africanês? Esse questionamento será reintroduzido nas discussões acerca dos aspectos metodológicos e de abordagens práticas.

Conforme a leitura de Maculan e Lima, em 2014, e Nunes, Maculan e Almeida, Maurício Barcellos, em 2021, constata-se a função estrutural das relações semânticas no sistema conceitual, das quais requer uma análise acurada de unidades verbais e suas especificações no tocante das valências verbais e no valor da alteração de significação de termos compostos ou sintagmas. Esse assunto implica na qualidade da rede que conecta os conceitos, principalmente, relativos às denominações de ações e de processos, bem diferentes daquelas unidades substantivas.

Com Souza e Pestana em 2017, as relações semânticas são visualizadas como elos na estruturação de tesouros, ontologias e taxonomias, cujo foco central desse entendimento

buscou reintegrar aquilo que expressa as noções de conceito e termo a partir das teorias da Classificação, Terminologia Geral e do Conceito.

Desde as investigações apontadas no referencial, o conceito de **comunidade discursiva** (*discourse communities*) se mostrou parte integrante de toda a narrativa teórica e metodológica da AD, estando presente nas definições e na própria noção de domínio. Analisando Hjørland e Albrechtsen em 1995, percebe-se que o estudo de domínios do conhecimento se articula no estudo das comunidades de discurso ou comunidades de pensamento, em determinados trechos os autores colocam esses conceitos como equivalentes e semanticamente semelhantes, apesar de não detalharem a respeito. Nota-se, ainda, “comunidade de pensamento” como sinônimo ou grupo detentor de um discurso ou de um pensamento (representação de um determinado conhecimento).

Amorim e Café concordaram com os apontamentos de Hjørland e afirma que a comunidade discursiva é aquela que

[...] se constitui a partir dos discursos e que é determinante das ferramentas, das linguagens, dos conceitos, das estruturas de informação e das necessidades de informação e dos critérios de relevância num domínio. Nossa proposta está de acordo com a composição heterogênea que Hjørland **confere à comunidade discursiva, na medida em que afirma que ela compreende atores, canais de informação, instituições, serviços de informação e relações sociais que envolvem tais atores**. Contudo, Hjørland não direciona especificamente como e em que o analista de domínio deve considerar tais componentes, sobretudo os agentes não-humanos. Nesse sentido, acrescentamos a importância de identificar os componentes não-humanos e explicar sua interação na produção da comunidade discursiva (Amorim; Café, 2017, p. 83).

Amorim e Café apostaram que, para além da identificação de atores e demais grupos, esta comunidade de discurso pode ser estendida para outros agentes não humanos (fatores, elementos, objetos, entre outros componentes que interagem com agentes humanos). Nesse interim, a condição de agente humano detentor do discurso está ligada aos protagonistas e aos usuários informacionais que comunicam e constroem os alicerces dos seus discursos em virtude de serem amplamente disseminados e tidos como relevantes.

Nascimento e Marteleto (2008) ao objetivarem a exploração do fenômeno da informação em determinado campo social e domínio – por meio da prática informacional e modo de agir que dá identidade a um grupo – reconheceram que as comunidades discursivas não são entidades autônomas, contudo, são construções sociais distintas e formalizadas por indivíduos cujo pensamento, a linguagem e o conhecimento atingem o ponto de sincronia e nutrem tal coletividade. Além das autoras recorrerem a teoria do discurso de Swales e ao

conceito de comunidade de Hjørland, igualmente, sustentam suas argumentações na noção de campo e das estruturas constitutivas do espaço social Bourdieu.

Hjørland e Albrechtsen, Hanne em 1995 e Hjørland em 2017, trataram do conceito de comunidade discursiva não somente pela formalidade dos sujeitos sociais na convergência dos discursos ambientados e característicos de seus interesses, mas percebem essa em decorrência da divisão do trabalho e dos fatores e condições que moldam possíveis relações – seja no ambiente trabalho e pela sucessão de propósitos econômicos e políticos – das quais tornam transitável o elo comunicacional entre o sujeito e a unidade de domínio (corpo de conhecimento compartilhado por outros sujeitos) que congregam na comunidade.

Wang e Qiu, em 2021 durante uma revisão acerca da Análise de Domínio, convergem, em partes, para alguns tópicos explanados por Nascimento e Marteleto, principalmente, aquelas direcionadas à estrutura social e o seu entrelaçamento com o conceito de comunidade discursiva. No entanto, Wang e Qiu se encaminham para a noção apoiada na epistemologia comunitária:

A epistemologia comunitária afirma que, porque cada comunidade discursiva assume um papel social diferente determinado pela divisão social do trabalho, ela desenvolve os seus paradigmas, regras, convenções, normas e rotinas, que são explícitas, formalizadas ou tácitas, não codificadas. Diferentes comunidades discursivas têm diversos interesses sociais, estruturas cognitivas coletivas e sistemas discursivos, respectivamente. Este facto determina que as comunidades darão interpretações diferentes quando se depararem com a mesma informação e adoptarem o mesmo sistema de informação. Assim a informação terá significados diferentes e o sistema de informação produzirá funções diferentes para cada comunidade (Wang; Qiu, 2021, p. 1030, tradução nossa, grifo nosso).

Aqui, os autores explicaram a constituição dessas comunidades. Nesse trecho é importante considerar, recordando a visão de campo de Bourdieu e da Teoria do Discurso, a existência das instituições e conjecturas que modificam os espaços, o modo de interação e as implicações na formação intelectual e profissional dos sujeitos. Essas instituições e conjecturas detêm influência nas aproximações, nos afastamentos e nas ações dos sujeitos, até mesmo aqueles que participam de uma comunidade de discurso. Um exemplo claro é o das organizações responsáveis pelo desenvolvimento e financiamento de pesquisas e produtos farmacêuticos, nutrindo, assim, grupos de pesquisadores (aglomeradas em comunidades discursivas) em meio a divulgação de periódicos científicos e de anúncios publicitários destinados às mídias televisivas e as demais.

O não sucesso dos sistemas de classificação, linguagens documentárias, e outros instrumentos e serviços direcionados aos públicos informacionais, implicam em **Viés**, ou do inglês **Bias**, enquanto manifestação de preconceitos e conflitos para determinados grupos e

sujeitos. Como forma de contornar essas tendências, surgem propostas ancoradas em abordagens e fundamentos que buscam minimizar e mitigar tais problemáticas. Nesse intuito, Feinberg (2007) expõe uma crítica sobre a responsabilidade das Bias e a apresentação da criação de um sistema de informação situado no pensamento de Donna Haraway. Com a autora percebeu-se que os sistemas idealizados para comunidades discursivas, de acordo com a abordagem analítica de domínio descrita por Hjørland e Albrechtsen (1995), podem ser visualizados como resposta para tais problemas.

O argumento de Feinberg, em 2007, está na questão do foco da Análise de Domínio de Hjørland se concentra em um único domínio por área temática e que nem sempre a base da análise em opiniões de especialistas condiz com o reconhecimento de determinados preconceitos. Ou seja, há estratégias que precisam ser definidas até mesmo para identificar domínio que se moldam em contexto multidisciplinar. Não o bastante, após os argumentos, podemos entender que a autora chega à conclusão de que o preconceito permanece com a abordagem analítica de domínio e, em decorrência disso, propõe a introdução do design de sistemas de informação. Esse design se volta para as particularidades e pode tornar claros os indícios das Bias e permitir alternativas adequadas para reverter tais posicionamentos.

Milani e Guimarães, em 2010, apresentaram os vieses de modo a reconhecê-los como parte integrante das questões fundamentais de elaboração das linguagens de indexação, demonstrando abordagens teóricas acerca das demandas feministas e suas responsabilidades diante do predomínio de vertentes machistas e tendências misóginas. O conceito de viés em linhas gerais aponta para ampliar as ações de acessibilidade e acesso, requisitando, em muitos casos, uma readaptação de métodos e princípios que absorvam outras vertentes epistemológicas não exploradas ou se quer introduzida na criação de sistemas de informação e organização.

É imprescindível destacar que Milani e Guimarães, em 2010, além de propor uma categorização para identificar preconceitos relativos às questões de gênero, dialogaram com o trabalho de Hope Olson de 2007, *How we construct subjects: a feminist analysis*, quando ambos percebem que os nossos sistemas de classificações podem apresentar características relevantes para a verificação dessas inadequações no uso de termos e suas relações paradigmáticas e sintagmáticas. Esses autores não tratam em específico de uma vertente do feminismo voltado para comunidades étnicas, por exemplo, mas buscam em linhas gerais prosseguir e promover um diálogo profícuo.

Os conceitos imbricados ao termo garantia estiveram marcantes na CI quando o assunto é aquele das abordagens que auxiliam na prática dos sistemas de organização e nas linguagens utilizadas para indexação. A literatura pontua e difere algumas dessas garantias. A

primeira a receber destaque, com mais ênfase nos artigos recuperados acerca da Análise de Domínio, é a **garantia literária** ou *literary warrant*.

Beghtol em 1995 com “*Domain analysis, literary warrant, and consensus: the case of fiction studies*” expõe, de forma técnica, a utilização desse conceito segundo o seu fundador, Hulme. Essa se qualifica como suporte para um estudo de AD que subsidia a quantificação de assuntos indexados na bibliografia da *Modern Language Association*. Na leitura, é entendido que o conceito de garantia literária é a solução de uma mensuração de assuntos e inferências representantes de uma literatura, por sua vez, essa respalda a utilização de termos em sistemas de classificações e linguagens de indexação. Entretanto, o direcionamento dado recorda ainda a diferenciação entre o consenso, pois ambos os conceitos têm suas existências e funções no desenvolvimento de sistemas:

Hulme (1911-19 12) cunhou o termo “garantia literária” para descrever a lógica que ele preferia para definir e nomear classes em um sistema de classificação bibliográfica, e o conceito de **garantia literária tornou-se fundamental nas discussões sobre acesso a assuntos** (por exemplo, Fraser, 1978; Rodríguez, 1984) [...]. Bliss (1939) usou o termo “consenso” para as opiniões que os cientistas (amplamente concebidos para incluir os humanistas) e os educadores partilham sobre as estruturas das suas disciplinas. Bliss sustentou que o consenso deveria estar subjacente aos sistemas de classificação bibliográfica; mas, ao contrário do conceito de garantia literária, o consenso não tem sido amplamente explorado nas discussões sobre o acesso aos assuntos. Na prática, contudo, o consenso tem sido geralmente considerado uma parte necessária do desenvolvimento e da revisão dos sistemas de classificação [...]. A **garantia literária** e o **consenso** não são conceitos mutuamente exclusivos ou opostos, e as relações entre os dois não foram totalmente explicadas na literatura. Para desenvolver uma fundamentação para um estudo quantitativo do conteúdo da literatura dos estudos de ficção, foi útil explorar algumas relações pertinentes entre eles (Beghtol, 1995, p. 31-32, tradução nossa, grifo nosso).

Alvarenga e Dias em 2012, assim como Beghtol, se apoiaram no conceito de Hulme e direcionam sua pesquisa na perspectiva da AD. Contudo essas últimas autoras apresentam o estudo de domínio voltados para o cenário da gestão arquivística ao utilizar a garantia literária para alicerçar a prática em organizações que têm suas atividades essenciais firmadas, exclusivamente, em documentos. Além disso, destacam outras garantias: literatura/documental, cultural, do usuário, organizacional, terminológica (Alvarenga; Dias, 2012).

Em outro percurso, Coladangelo (2020) confirmou essa garantia no desenvolvimento de tesouros, mais precisamente, buscando alicerçar uma metodologia de estruturação de termos do domínio da dança folclórica. Esse estudo esteve direcionado para o uso em modelos e em organização do patrimônio cultural, além disso resgata traços e costumes indígenas do qual o autor buscou retratar e eleger como o domínio.

O destaque dado a garantia literária em relação as demais garantias, é conferido na revisão crítica escrita por Barité, Fernandez-Molina, Guimaraes e Moraes (2010) ao sugerirem que essa garantia pode ser aplicada em três âmbitos: no contexto que possa justificar as relações presentes em um sistema de organização do conhecimento; no respaldo para elaboração de mapas de domínios; e, por último, como dispositivo de avaliação de estruturas conceituais.

Ambas as garantias, expressam uma decisão de optar por um caminho que ligue um usuário de um sistema até a informação desejada. Ou seja, cada uma é a prova de que um assunto é tratado, um termo é consolidado, um conceito é entendido ou qualquer unidade de conhecimento (modelo, teoria etc.) é amplamente utilizado pelos sujeitos em um determinado contexto.

Quanto a **garantia cultural**, Alvarenga e Dias (2012) e Oliveira e Araújo (2012) concordaram e recordaram a definição de Beghtol, destacando ser argumentações tidas por considerar que as classificações e as relações semânticas dependem de um contexto cultural para estabelecer seu significado e função. Alvarenga e Dias em 2012 e Rauch e Barité em 2020 são responsáveis por tratar e visualizar esse conceito na prática, ambos os trabalhos foram recuperados e identificados como estudos em Análise de Domínio. Moura e Assis em 2010, Gomes e Frota em 2019 e Coladangelo em 2021 consolidam diálogos que permitem reconhecer a questão cultural atrelada às necessidades de introduzir novas abordagens e perspectivas para o desenvolvimento de sistemas de organização do conhecimento, como vocabulários especializados e tesouros.

Para compreender a garantia cultural, Rauch e Barité (2020) apresentaram quatro elementos teóricos descobertos no contexto da Organização do conhecimento, são eles: a necessidade de adequação do conceito de cultura; o direcionamento dado às dimensões locais do conhecimento; o intuito de promover a classificação tendenciosas para favorecer de grupos tratado como socialmente minorizados e relegados; e a questão ética expressa no respeito à integridade do pensamento cultural de uma dada comunidade.

A **garantia de uso**, ou **garantia do usuário**, é outra noção bastante utilizada para permitir o protagonismo dos sujeitos em relação às outras garantias aqui ressaltadas. Podemos conferir Moura em 2009 e Moura e Assis em 2010, que essa garantia é mencionada e acompanhada com outras garantias (cultural e literária), não sendo mencionada ou tratada de forma isolada.

Gomes e Frota, em 2019, exploraram a perspectiva social na construção dos tesouros ao integrar garantias como a cultural e a filosófica. A **garantia filosófica** pode ser assimilada enquanto a influência admitida na utilização da literatura científica como primeira a

respaldar os processos de utilização e elaboração de vocabulários e linguagens, pois essa agrega e divulga os construtos filosóficos, principalmente, aqueles que desvendam os aspectos ontológicos e seu papel na atribuição de redes, modelos e relações conceituais.

As garantias estrutural e semântica foram detectadas em Moura, 2006, e Gomes e Frota, 2019, embora suas definições conceituais não foram detalhadas. Oliveira e Araújo (2012) citam Moura e Moreira (2006), esclarecendo que a **garantia estrutural** é aquela que encontra justificativa na estrutura do tesouro, ou seja, o esquema e as estratégias estabelecidas pelas relações e ordenações dos termos são utilizados como guia e argumento para tal.

Em paralelos aos artigos relativos às Bias, às garantias, às discussões acerca da adequação dos sistemas organização do conhecimento e às variedades de grupos, encontrou-se os conceitos de **hospitalidade cultural** e **interseccionalidade**. Nessa perspectiva, Coladangelo, em 2021 durante a descrição de sua proposta do uso de anotação semântica de anéis de sinônimos, assinalou a importância das questões culturais e a função do princípio de hospitalidade cultural no desenvolvimento de um sistema próprio do contexto cultural e a partir da participação do usuário. Esse trabalho utilizou a Análise de Domínio nos resultados preliminares de descoberta da temática escolhida e como forma de prospecção das etapas tidas no processo de anotação semântica.

A interseccionalidade é entendida como ferramenta analítica que permite um enquadramento, colocando a identidade individual no eixo de muitas outras identidades sociais e políticas (Coladangelo, 2021). Ou seja, esse conceito pode ser percebido como um prisma que visualiza um ou mais aspectos simultâneos e componentes da identidade de um sujeito, atestando o reconhecimento da pluralidade de identidades sociais – gêneros, raças e etnias – conforme a importância dada as suas características, suas linguagens, suas manifestações e seus direitos civis. Ressalta-se que Coladangelo buscou introduzir esse conceito como um guia para alcançar princípios éticos na sua prática de organização do conhecimento.

Com Barité e Rauch em 2020, a hospitalidade cultural é mencionada na discussão sobre a garantia cultural. Apesar de Coladangelo, em 2021, dar notoriedade à hospitalidade cultural, a definição desse conceito não é detalhada. Porém, a forma como esse princípio é informado durante o processo de identificação para candidatos aos termos de um vocabulário deixa explicitado que sua noção integra elementos culturais. Isso esteve evidenciado na modelagem semântica, e na análise interseccional, quando esse autor buscou apontar variáveis essenciais para alcançar sua proposta, dentre elas a etnicidade, a raça, o gênero, a linguagem.

O conceito de **cultura** foi tratado por Rodrigues em 2012 e 2015, além de ser mencionado em Rauch e Barité em 2020, ambos indexados em Análise de Domínio. Percebeu-

se uma necessidade dos autores de resgatar esse conceito nas discussões teóricas da organização do conhecimento, associando sua complexidade à recorrência dos vários sentidos e significados e à influência das várias escolas de pensamento e disciplinas.

Para Rauch e Barité (2020) o conceito de cultura, assim como o conceito de garantia cultural, não são neutros, pois privilegiam determinadas formas de organização do conhecimento, já que substituem os critérios de objetividade e neutralidade por aqueles de caráter cultural e ligados aos valores de uma comunidade.

Rodrigues (2015) apontou que o termo cultura tem origem etimológica da palavra latina *cultura* e seu caráter é complexidade e possui inexorável ambiguidade. Para elucidar conceitualmente, o autor utilizou a Análise de domínio para descobrir o domínio da cultura. A análise e os apontamentos históricos e teóricos recorrem aos autores da sociologia, da antropologia e das teorias que cercam tal ideia, dentre eles Zygmunt Bauman e Claude Lévi-Strauss. Algumas definições readmitem as principais proposições das teorias sociais e antropológicas, adentrando em perspectivas que confinam nas representações humanas.

Para Loureiro (2001), a cultura é entendida como uma configuração intelectual, artística e moral de um povo ou, mais amplamente, de uma civilização, e que pode ser compreendida no processo de seu desenvolvimento histórico ou num período delimitado de sua história. Para Nunes (2004, p. 9) é a “soma de todas as criações que melhoram a condição material dos homens ou que expressam a vida intelectual e moral”. A cultura reproduz o resultado do esforço de autoafirmação da espécie humana, que se desprende da condição da animalidade até atingir, passando pela selvageria e pela barbárie, os mais altos graus de desenvolvimento. Essa ideia, no entanto, é incompleta e foi corrigida e ampliada pelo conceito antropológico de cultura. A cultura, como termo geral, significa a herança social total da humanidade. No conceito antropológico, a cultura é, para cada agrupamento humano, um conjunto de modos de proceder e pensar segundo estruturas normativas variáveis e particulares que sustentam diferentes padrões de pensamento e ação. A Antropologia e outras áreas de estudo ensinam a olhar e a compreender a variedade das culturas, indicando a síntese da criação do universo humano por meio de normas, indissociável da linguagem, técnicas, conhecimento, religião, valores éticos, estéticos e políticos existentes no conjunto de uma organização social determinada. Aranha e Martins (1993) comentam que a palavra cultura tem vários significados, mas no campo da Antropologia a cultura é o que o homem produz ao construir sua existência, as práticas, as teorias, as instituições, os valores materiais e espirituais (Rodrigues, 2012, p. 12, grifo nosso).

O olhar antropológico incumbe a diversidade e a natureza plural das manifestações que desenha o conceito de cultura. Embora ao revisitar esse conceito, alguns questionamentos e apontamentos críticos devem guiar o pesquisador ou leitor, quando a direção dada para “cultural” não seja aquela que admite determinadas culturas e menospreze outras, das quais, em muitos casos, não são tratadas como cultura. Isso pode ser visto com a aplicação dessa noção – noção envolvida pelas faces da moral, do intelecto e do artístico – para populações periféricas,

determinados grupos étnicos e marginalizados, pois seus sujeitos podem ser vítimas das tendências de segregação, extermínio e violência. Essas tendências podem ser facilmente visualizadas na construção de vocabulários estruturados ou classificações bibliográficas, por consequência, a temática cultura se estabeleceu em paralelo as noções de garantia cultural, etnia, raça, gênero e interseccionalidade.

Em um pensamento mais cristalizado no âmbito organizacional e distinção das culturas na macroestrutura global, Wang (2019), apontou duas definições: a primeira da *Organizational Behavior Effectiveness* (GLOBE) ao definir cultura como conjunto de interesses, valores, crenças, identidades, interpretações ou significados compartilhados em eventos significativos, estando enraizados em experiências comuns entre membros e transmitidos através de gerações; e a definição do próprio autor ao tratar cultura enquanto uma construção coletiva que diferencia algumas pessoas de outras, além de ser uma influência comportamental.

Moraes (2018) tratou da visão homogeneizadora, que está circunscrita na ideia de cultura perante o contexto brasileiro, ao citar a Carta do Folclore Brasileiro e seus desdobramentos. Essa análise sucinta a questão envolvendo o conhecimento popular e sua relação entre o saber do cientista e saber do não cientista (pessoa com formação baseada em ensino médio ou não frequentador de instituições acadêmica-científica). Isso impacta diretamente em questões particulares dos sistemas de organização do conhecimento, como a alocação desse termo e conceito em teorias sociais cujas conjecturas estão na base dos saberes populares, religiosos ou culturais.

Ainda em Moraes (2018), o debate da tensão identitária na organização do conhecimento é ilustrado com auxílio da abordagem epistemográfica, cuja relação está estreita com o conceito de **etnoconhecimento** e sua representação. O estudo da autora foi recuperado nos documentos referentes à Análise de Domínio e aos trabalhos acerca dos tesouros, enfatizando uma temática convergente e associada a ambos os assuntos. A discussão foi ambientada nos estudos da Organização e Representação do Conhecimento e das críticas direcionadas as epistemologias tradicionais. A introdução do conceito, segundo Moraes (2018), perpassa por Garcia Gutierrez em 2006 – acerca da perspectiva transdisciplinar como componente da organização horizontal e interativa dos conhecimentos; Miranda em 2007 e 2012 – com a abordagem epistemográfica; Begthol em 1995 – na exposição da garantia a partir do domínio; Lopes Huertas e Ramírez – sob o olhar da terminologia de gênero; e Franca e Silveira em 2014 - a temática da representação bibliográfica do conhecimento indígena registrado. A definição da autora explicou no reconhecimento da função desta categoria:

O etnoconhecimento é uma categoria que melhor espelha a diversidade étnica e racial, e está intimamente relacionado à realidade social de um determinado grupo e também à oralidade como forma de transmissão de saberes produzidos coletivamente. No bojo da abordagem epistemográfica, o conflito identitário é aqui vislumbrado a partir da identificação de estratégias de poder e de resistência dos excluídos do discurso hegemônico de um domínio. No caso do TEGM, a luta das mulheres negras no campo do feminismo, pela priorização de termos e conceitos que representam as especificidades da mulher negra e que não constam do discurso hegemônico feminista brasileiro da década de 90, dá visibilidade a uma prática identitária excludente. Também o Tesouro de Folclore e Cultura Popular, elaborado sob os auspícios do Centro Nacional de Folclore e Cultura Popular, apresentou um discurso hegemônico no campo folclorista. Focamos como este instrumento documentário espelhou um discurso que deu prioridade à interferência ativa do Estado Brasileiro na interpretação e no fomento do folclore, resultando na construção de uma linguagem hegemônica que concebia o termo folclore como sinônimo de cultura popular brasileira e ignorou a diversidade étnica e racial. A reflexão aqui colocada sugere a abordagem epistemográfica como uma das estratégias (Moraes, 2018, p. 1).

É perceptível reconhecer que esse conceito submerge dos encontros estabelecidos nas epistemologias ancoradas no colonialismo e no universalismo da ciência, pois se refere às tradições orais, às representações e às manifestações do pensar e do agir dos povos tradicionais, sejam presentes na literatura, nas histórias orais, nas artes, na culinária e outros.

A reflexão diante do etnoconhecimento em Moraes, percebe-se, por exemplo, que a solução para o emprego de determinados termos em tesouros direcionados às temáticas indígenas está em introduzir fundamentação teórica e metodológica direcionada às epistemologias e ao conhecimento produzido e mantido pelos grupos e povos indígenas. Isso vale para qualquer conhecimento de substrato étnico-racial ao considerar a localização, as fronteiras e as identidades enraizadas em tais saberes.

O etnoconhecimento é uma categoria localizada no mesmo caminho crítico elucidado pelas questões étnico-raciais promovidas a partir da apreciação dos conceitos do pensador Aníbal Quijano. No artigo “Por uma reflexão sobre a organização e representação de conceitos decoloniais na América Latina: O pensamento de Aníbal Quijano à luz da Análise de Domínio” de 2023, Albuquerque apresentou os seguintes conceitos importantes dentre os estudos aqui citados, são eles: Modernidade/Colonialidade, Decolonialidade, Colonialidade do Poder, Classificação social (raça e gênero).

Albuquerque (2023) descreveu o conceito de **colonialidade do poder** de Quijano e de Ballestrin. Para Ballestrin (2013 apud Albuquerque, 2023) a colonialidade do poder se apresentou como uma configuração do processo estrutural do sistema-mundo moderno/colonial ao permitir articulações de lugares periféricos, contextualizados na divisão internacional do trabalho e no estabelecimento das hierarquias agenciadas pelas questões étnico-raciais globais.

Outro ponto de destaque: Albuquerque (2023) apontou que esse conceito não está somente ligado aos âmbitos político e econômico, mas está ramificado para outros âmbitos. Percebeu-se, neste ponto, que os referidos âmbitos não são direcionados somente ao entendimento do discurso formado na base da América Latina ou nos estudos da CI com um papel científico, empírico e teórico local. Pelo contrário, esse pensamento resgata a construção teórica e conceitual com base em elementos e problemas globais.

A nova configuração desse poder colonial tem suas consequências na segmentação social. A partir daí o conceito de **classificação social**, na obra de Quijano, para Albuquerque (2023, p. 15) “é derivado do entendimento que as diferenças raciais são espécies de métodos, que distribuem os sujeitos em seus determinados papéis na estrutura social do trabalho”. Conforme essas diferenças, são traçados os problemas sociais recorrentes das discriminações (classe, raça e etnia) articuladas em estratégias de poder presentes em diversos segmentos, principalmente, aqueles dos grupos que detêm o poder intelectual tecno-científico e político.

A ideia de **raça**, discutida no referencial teórico com base nas teorias da etnicidade com Poutignat e Streiff-Fenart (1998) e em autores como Munanga (2004) e Raffestin (1993), se revelaram, da mesma forma, na centralidade das proposições e diante da característica de poder e de estabelecimento das nações e grupos étnicos. Assim, Albuquerque apontou para a interseção de diálogos de outros conceitos, como gênero e dominação:

A ideia de raça é um dos elementos estruturais de dominação colonial eurocêntrica e está, junto a dominação pelo sexo, em uma sociedade patriarcal, dominação da mulher negra, do homem negro, dos indígenas e de minorias étnicas na visão eurocêntrica, inserida no processo de continuidade da subjugação dos povos pelo colonialismo. Quijano (2000) aloca raça, gênero e trabalho como elementos principais que firmam o modelo de capitalismo colonial/moderno. Desta forma, tem-se os conceitos em uma relação histórica: a exploração, dominação e conflito ocasionados pela perpetuação dos processos coloniais são uma marca produzida pela colonialidade do poder, que traz em seu cerne a estruturação hierárquica do sistema-mundo. Assim, a classificação social é produzida e reproduzida pela colonialidade do poder e serve de estrutura elementar, para a apreensão subjetiva e espacial dos diversos aspectos dos povos da América Latina (Albuquerque, 2023, p. 15).

No trecho acima, apesar das questões étnicas-raciais terem características e descrições distintas conforme o local e cultura, as menções e as discussões desses conceitos, principalmente, os étnicos e raciais, reverberam o incomodo causado pela ausência desses temas no plano de fundo da CI e da Organização do Conhecimento (*Knowledge Organization*), seja em caminhos epistêmicos e teóricos, ou até mesmo no simples fato de associa-los, em específico, aos artigos e palestras com o foco nos usuários e nas classificações presentes nos

âmbitos sociais (classificação do trabalho, classes sociais, comunidades ou partidos políticos). Albuquerque ainda menciona o conceito de gênero nesse estudo de 2023.

Santos, Costa, Barros e Vital, em 2020, apontaram os termos **negro** e **afro-brasileiro** no contexto tanto da literatura recuperada em tesouros, quanto as publicações que abordam questões de raça, etnia, cultura e tesouros. Os autores ressaltaram a importância de reconhecer a representação terminológica sobre a população negra. Por outro lado, são com base nessas menções desses conceitos que demonstraram uma preocupação de citar e envolver cada vez mais os conceitos de negritude e de outros grupos étnico-raciais. O estudo não descreveu as características e os apontamentos históricos relativos aos termos negro e afro-brasileiro.

O cenário de lutas do feminismo possibilitou refletir tanto sobre os aspectos conceituais hibridadas na ideia de **gênero** – já que a classes sociais e outras categorias implicam no sujeito social- quanto a elucidação de outras vertentes que viriam a representar outros grupos, como o feminismo negro brasileiro.

Nessa apreensão conceitual, são destacaram definições mencionadas Ribeiro, Decourt, e Almeida em 2017, retratando o estudo do domínio de gênero e sua utilização na construção de linguagens documentárias através do mapeamento de termos. Contudo, as citações conceituais elencadas pelos autores, apontou gênero enquanto construção social e sob a forma do binômio homem e mulher. Reconhecendo uma ausência, escassez ou silenciamento de outras noções que contemplam a diversidade de gênero e a inclusão das pessoas intersexo.

Sousa e Tolentino em 2017, observou na história a questão de dominação e reafirmação dos papéis na sociedade como alguns pontos a serem analisados frente à noção de gênero. Ademais, os autores retratam esse assunto a partir da crítica aos termos e palavras machistas e possibilitaram reintroduzir perspectivas de igualdade de gênero e conduzirem um destaque para pontos de vistas do feminismo na construção de descritores nos tesouros e outros vocabulários. O mesmo pode ser conferido no artigo de López-Huertas em 2007 recuperado na temática sobre léxicos, a autora fundamentou sua pesquisa envolvendo domínios interdisciplinares voltados para a recuperação da informação em temáticas de gênero, cujo propósito está no tratamento de léxicos utilizados na indexação.

Moraes em 2018, demonstrou preocupação quanto à dimensão das questões raciais e de gênero, relatando pontos imprescindíveis para compreender os esforços de grupos que sofrem constantemente com os estigmas da divisão racial e sexual no trabalho, como exemplo as feministas negras. Nessa perspectiva, observou-se que os autores latinos, e dos poucos

africanos e asiáticos representam mais a temática étnico-racial do que os autores norte-americanos e do leste europeu.

Pinho e Guimarães, em 2012, apresentam um estudo acerca das inconsistências terminológicas de termos e assuntos relacionados a **homossexualidade masculina** na indexação. A pesquisa utilizou a AD para mapear esses indícios e explorar as questões éticas direcionada a essa comunidade no contexto brasileiro. Esse tipo de estudo demonstra a carência na compreensão de termos e conceitos em sua relação direta com as pessoas dessa comunidade tão plural, para então eliminar ou mitigar os efeitos do preconceito e recorrer às representações estabelecidas de seus membros.

Olson em 2007 no capítulo intitulado de “*How We Construct Subjects: A Feminist Analysis*” destacou os questionamentos acerca das questões de gênero nas estruturas classificatórias dos esquemas de classificação e na prática de desenvolvimento dos assuntos. A autora procurou suscitar até mesmo nas relações semânticas, que a prática classificatória e de organização tradicionais perpassa por fundamentos que excluem uma análise feminista. Apesar de suas críticas serem bem diretas no desenvolvimento desses esquemas e assuntos, Olson reconhecer que as influências, como a lógica formal, já se qualificam naturalmente como idealizadoras e mantenedoras das representações típicas do gênero masculino. Alguns exemplos são citados e o tecer das críticas foram encaminhados para as estruturas de relacionamentos de tesouros e vocabulários:

As relações definidas pelo homem, tanto paradigmáticas como sintagmáticas, podem ser ideais, mas algumas técnicas automáticas semelhantes às funções de encontrar mais-como-este podem ser desenvolvidas em um nível mais sofisticado, talvez com processamento de linguagem natural. Mapear as complexidades dos relacionamentos acrescentaria maior conectividade, mesmo que esses relacionamentos não fossem nomeados. Tal abordagem poderia levar ao desenvolvimento de normas baseadas na lógica indutiva e não na lógica dedutiva (Olson, 2007, p. 558, tradução nossa).

Desse modo, a autora destacou algumas considerações construtivas, dentre elas referentes às relações mencionadas acima nessa seção e que requer alterações metodológicas, como por exemplo a substituição da abordagem indutiva por outras abordagens mais inclusivas.

A sigla **LGBTQI –lésbica, gay, bissexual, transgênero, queer e intersexo** - foi tratada no artigo de Golub, Bergenmar e Humelsjo em 2022, englobando esses termos que expressam e representam algumas das categorias da **diversidade sexual e de gênero**. A introdução desses termos eleva a temática para o plano de fundo dos trabalhos realizados em conjunto com esses grupos. Como exemplo, Golub, Bergenmar e Humelsjo direcionaram a temática dos tesouros para a literatura sueca lésbica, gay, bissexual, transgênero, queer e

intersexo, cuja importância dessa pesquisa empírica esteve em abrir uma interfase de discussão para colocar em destaque o protagonismo das comunidades de diversidade sexual, principalmente, a luz das teorias de gêneros e das sexualidades, como o caso da *Teoria Queer*. Nesse ponto, ao direcionarem perguntas para usuários de livros de ficção LGBTQI, o trabalho fomentou uma certa proximidade dos estudos de usuários, da terminologia de especialidade ou léxicos e vocabulários dessa temática para compreensão dos conceitos dentro desse universo.

Na mesma direção da estruturação de termos em tesouros, Bissoli, Covello, Pisseli, e Santos em 2018, apontaram uma sigla atualizada que inclui um maior número de representações da diversidade sexual e de gênero, a sigla LGBTQIAP+ (lésbicas, gays, bissexuais, travestis, transexuais, transgêneros, queers, intersexuais, assexuais e as demais orientações e identidades). Essa sigla além de representa a diversidade, também, reforça a noção de inclusão e de direitos civis dessa comunidade.

6.2 Aspectos teóricos

O mapeamento do referencial teórico esteve com expressividade nos artigos que mencionam a Organização do Conhecimento e os alicerces da CI. A principal problemática cercou as imprecisões quanto à noção de paradigma e seus limites formais dentro do próprio discurso dominante na OC. Retoma-se aqui a noção de paradigma de Kuhn para auxiliar nas bases epistêmicas. Vejamos no Quadro 8 os autores identificados e suas contribuições:

Quadro 8 – Teorias, paradigmas e epistemologias

Teorias, paradigmas e epistemologias	Análise de Domínio	Tesauros	Léxico	Etnia, Raça, Cultura e Análise de domínio	Tesauros e Análise de Domínio	Etnia, Raça, Cultura e Tesauros
Teoria do Conceito	Martines, Alexandre Robson; Almeida, Carlos Cândido de - 2021	Gomes, Hagar Espanha; Campos, Maria Luiza Almeida - 2004 Campos, Maria Luiza Almeida; Gomes, Hagar Espanha - 2005 Hjørland, Birger - 2009 Stock, Wolfgang G. - 2010 Carlan, Eliana; Bräscher, Marisa - 2011 Souza, Jacqueline A.; Pestana, Olívia - 2017 Gomes, Hagar Espanha - 2021 Ferreira, Ana Carolina; Maculan, Benildes Coura Moreira dos Santos – 2020	Martines, Alexandre Robson; Almeida, Carlos Cândido de - 2021			
Teoria da Classificação Facetada	Rodrigues, Anderson Luiz Cardoso - 2012; 2015	Campos, Maria Luiza Almeida; Gomes, Hagar Espanha - 2005 Carlan, Eliana; Bräscher, Marisa - 2011 Souza, Jacqueline A.; Pestana, Olívia - 2017 Ferreira, Ana Carolina; Maculan, Benildes Coura Moreira dos Santos - 2020 Gomes, Hagar Espanha – 2021		Rodrigues, Anderson Luiz Cardoso - 2012; 2015		
Semiótica de Peirce	Reis, Daniela Majorie Akama dos; Fujita, Mariângela Spotti Lopes; Santos, Plácida Leopoldina Ventura Amorim da Costa; Zafalon, Zaira Regina - 2018 Martines, Alexandre Robson; Almeida, Carlos Cândido de - 2021	Martines, Alexandre Robson; Moreira, Walter; Almeida, Carlos Cândido de – 2022	Martines, Alexandre Robson; Almeida, Carlos Cândido de - 2021			

Teorias, paradigmas e epistemologias	Análise de Domínio	Tesaurus	Léxico	Etnia, Raça, Cultura e Análise de domínio	Tesaurus e Análise de Domínio	Etnia, Raça, Cultura e Tesaurus
Semiologia de Saussure	Reis, Daniela Majorie Akama dos; Fujita, Mariângela Spotti Lopes; Santos, Plácida Leopoldina Ventura Amorim da Costa; Zafalon, Zaira Regina - 2018 Martines, Alexandre Robson; Almeida, Carlos Cândido de - 2021	Martines, Alexandre Robson; Moreira, Walter; Almeida, Carlos Cândido de – 2022	Martines, Alexandre Robson; Almeida, Carlos Cândido de - 2021			
Semiologia de Bakhtin	Reis, Daniela Majorie Akama dos; Fujita, Mariângela Spotti Lopes; Santos, Plácida Leopoldina Ventura Amorim da Costa; Zafalon, Zaira Regina - 2018 Martines, Alexandre Robson; Almeida, Carlos Cândido de - 2021	Martines, Alexandre Robson; Moreira, Walter; Almeida, Carlos Cândido de – 2022	Martines, Alexandre Robson; Almeida, Carlos Cândido de - 2021			
Empirismo	Hjørland, Birger - 2004b; 2005; 2022 Amorim, Igor Soares; Vianna, William Barbosa; Brasher, Marisa - 2019	Hjørland, Birger - 2009 Stock, Wolfgang G - 2010 Mostafa, Solange Puntel; Cruz, Denise Viuniski da Nova -2010				
Racionalismo	Hjørland, Birger - 2005; 2022 Amorim, Igor Soares; Vianna, William Barbosa; Brasher, Marisa - 2019	Hjørland, Birger - 2009 Stock, Wolfgang G – 2010				

Teorias, paradigmas e epistemologias	Análise de Domínio	Tesaurus	Léxico	Etnia, Raça, Cultura e Análise de domínio	Tesaurus e Análise de Domínio	Etnia, Raça, Cultura e Tesaurus
Realismo Epistemológico	Hjørland, Birger - 2004b; 2022 Thellefsen, Martin Muderspach - 2023					
Historicismo	Hjørland, Birger – 2004b; 2022	Hjørland, Birger - 2009 Stock, Wolfgang G – 2010				
Hermenêutica		Hjørland, Birger - 2009 Stock, Wolfgang G – 2010				
Positivismo	Hjørland, Birger - 2004b, 2005; 2022 Freitas, Juliana Lazzarotto; Bufrem, Leilah Santiago; Breda, Sônia Maria - 2016 Amorim, Igor Soares; Vianna, William Barbosa; Brasher, Marisa - 2019	Hjørland, Birger - 2009				
Pragmatismo	Hjørland, Birger – 2004b; 2022	Hjørland, Birger - 2009 Stock, Wolfgang G – 2010				
Teoria Geral da Terminologia		Campos, Maria Luiza Almeida; Gomes, Hagar Espanha - 2004 Carlan, Eliana; Bräscher, Marisa - 2011 Souza, Jacqueline A.; Pestana, Olívia - 2017 Gomes, Hagar Espanha – 2021				
Paradigma cognitivo na CI	Campos, Luiz Fernando de Barros; Venâncio, Ludmila Salomão - 2007 Silva, Jonathas Luiz Carvalho; Farias, Maria Giovanna Guedes - 2013 Rodrigues, Anderson Luiz			Campos, Luiz Fernando de Barros; Venâncio, Ludmila Salomão - 2006 Rodrigues, Anderson Luiz Cardoso - 2015		

	Cardoso - 2015 Thellefsen, Martin Muderspach – 2023					
Teorias, paradigmas e epistemologias	Análise de Domínio	Tesauros	Léxico	Etnia, Raça, Cultura e Análise de domínio	Tesauros e Análise de Domínio	Etnia, Raça, Cultura e Tesauros
Paradigma social na CI	Silva, Jonathas Luiz Carvalho; Farias, Maria Giovanna Guedes - 2013	Hjørland, Birger – 2009				
Sociolinguística	Moraes, Miriam Gontijo de - 2018	Moraes, Miriam Gontijo de -2018	León, David Leonardo García -2014 Guerrero, Gladys Yolanda Pasuy; Montoya, Claudia Liliana Agudelo - 2022	Moraes, Miriam Gontijo de - 2018		Moraes, Miriam Gontijo de - 2018

Fonte: Elaborado pela autora.

A ambientação e os diálogos promovidos por Silva e Guedes em 2013, resgataram os principais autores citados na literatura dos paradigmas da CI: Thomas Kuhn, Boaventura Santos, Capurro, Hjørland e Albrechtsen. A análise histórica e epistemológica elevou as interpretações do **paradigma cognitivo** em direção aos domínios e suas comunidades previsto na Análise de Domínio e a Hermenêutica da Informação, destacando o encontro com o paradigma social. A principal prerrogativa desse paradigma está em considerar o sujeito cognoscente como a entidade capaz de espelhar e guiar modelos mentais e comportamentais a serem seguidos no desenvolvimento de práticas, serviços e instrumentos informacionais.

O modelo de recuperação de informação de Belkin, Oddoy e Brooks – descritos em comunicações das décadas de 70 e 80 - analisado por Campos e Venâncio em 2007¹⁹, ilustrou um pensamento da abordagem cognitivista e seus impactos na recuperação da informação e, consequentemente, na representação do conhecimento, na representação da informação e na elaboração de linguagens e vocabulários controlados:

Representadas as necessidades de informação, bastaria, então, especificar os mecanismos e estratégias de recuperação, diversos para cada tipo de estado anômalo de conhecimento. Assim, os autores propõem um modelo de sistemas de recuperação interativo que determina as representações de estados de conhecimento anômalos,

¹⁹ A análise epistemológica e histórica dos autores acerca das abordagens físicas, cognitivistas e emergentes no campo da Ciência da Informação adotou o conceito de abordagem de Wersig em detrimento da concepção de paradigma.

tendo por base as necessidades de informação. O sistema constrói uma representação do conhecimento do usuário através de redes de associações entre palavras identificadas nas declarações narrativas do problema, examina essa estrutura, interpreta as características estruturais das representações para determinar os estados anômalos do conhecimento, e pesquisa as descrições de documentos estruturados com o objetivo de solucionar a anomalia. **O usuário, então, avalia os documentos recuperados e essa avaliação é utilizada, pelo sistema, para determinar as adequadas modificações nas representações das necessidades do usuário, assim como a troca da estratégia para a definição das características significativas do mecanismo de busca adotado** (Campos; Venâncio, 2007, p. 112, grifo nosso).

Logo, essa tendência encarrega a prática do tratamento descritivo e temático – catalogação e indexação – nos mesmos moldes cognitivistas da recuperação da informação. A implicação da estruturação dos assuntos e a composição dos sistemas de classificação eram alimentados nessa visão pelos experimentos e estudos comportamentais dos usuários, dos quais esses últimos se qualificam enquanto sujeitos que utilizam e participam diretamente dos sistemas de recuperação. Outras categorias e grupos não eram analisados coletivamente, além disso, a introdução de ferramentas e instrumentos tecnológicos denotavam uma analogia estabelecida entre o cérebro humano e as máquinas de processamento da informação, como os computadores.

Apesar de não mencionar o termo paradigma, Rodrigues (2015, p. 46) considerou o cognitivismo como influência na CI e afirmou que essa visão filosófica e epistêmica “[...] analisa o fluxo do processamento de sinais na mente sem observar o contexto sociocultural e a bagagem cultural de experiências vividas pelo mesmo indivíduo”.

Hjørland (2009), na descrição e análise da teoria do conceito, revelou que entre a década de 1980-1990, a visão cognitiva na CI esteve relacionada com a inteligência artificial e a psicologia cognitiva, em específico afirmou que naquela época, e em outros períodos anteriores, os conceitos eram estudados principalmente com base nos estudos experimentais dos processos cognitivos humanos. É possível ainda recordar as influências de outros paradigmas e abordagens cognitivas envolvendo pesquisas e experimentos na Linguística e na Ciência Computacional aplicada em processos de estruturação e mapeamento de linguagens.

Silva e Farias (2013) pontuaram, em um nível macro, a perspectiva da referência do paradigma cognitivo situado amplamente na CI e, em um nível micro, o estabelecimento desse referencial na interação das correntes teóricas influentes na CI²⁰, nesse último nível desenvolveram-se: o paradigma cognitivo behaviorista; o paradigma cognitivo funcionalista e o paradigma cognitivo estruturalista. Os autores reconhecem as críticas envolvendo a vertente cognitiva e a insuficiência em atender alguns questionamentos.

²⁰ Nesse trecho, percebeu-se que os paradigmas micro estabelecem o paradigma macro.

Nielsen (2004), relatando os fundamentos teóricos e históricos da construção de tesauros, esclareceu que as abordagens centradas no usuário e a perspectiva cognitiva receberam críticas por atenderem aos estudos de usuários individuais e separado, cada um em seu contexto, argumentando que as estruturas do conhecimento e o comportamento do usuário individual são moldados por meio da participação social em domínios. É nesse sentido que Silva e Farias em 2013, recorreram a Análise de Domínio e a Hermenêutica da Informação como duas propostas de estabelecimento do paradigma social (sociocognitivo) para corresponder aos problemas empíricos e teóricos não resolvidos pelo paradigma cognitivo. Thellefsen em 2023, conjuntamente, enfatizou que a Análise de Domínio é uma resposta a noção cognitiva.

Observando a literatura da tríade raça-etnia-cultura e Análise de domínio, Campos e Venâncio, anteriormente em 2006 com o artigo “o objeto de estudo da Ciência da Informação: a morte do indivíduo”, procuraram por meio de uma exposição epistemológica de metateorias questionar o objeto e o foco científicos da CI, indicando os possíveis impactos de algumas abordagens, princípios teóricos e a decadência de outras perspectivas. Nessa discussão, é argumentada a expressão máxima do individualismo ao adotar a orientação cognitivista ambientada na CI (Campos; Venâncio, 2009), fazendo refletir tanto na abordagem holística e interdisciplinar, quanto no sujeito frente aos elementos epistêmicos - a materialidade do documento e a natureza da informação, ademais suas implicações subjetivas e objetivas.

O ambiente desse debate é complexo, onde o cognitivismo é confrontado com a nova, ou as novas, configuração científica que identifica outras dimensões e elementos, principalmente, a do neopragmatismo e aquelas dos novos estudos que transferem o protagonismo da mente humana para outras realidades: aquelas dos sistemas e das máquinas com rapidez e funcionalidades no processamento das linguagens. Isso infere diretamente em outras formas de condução das práticas na CI, como aquelas que se valem do processamento de linguagem natural e aprendizado de máquina para realizar processos de tratamento da informação antes realizados por pessoas.

Diante desse contexto mencionado por Campos e Venâncio, pode-se dizer que é possível treinar um computador para selecionar termos indexadores de forma mais eficiente e rápida do que aquelas realizada na indexação semiautomática ou extração de Unitermos no período de estabelecimento das primeiras abordagens promissoras na recuperação da informação.

Analisando trabalhos que utilizam o paradigma cognitivo, foram encontradas duas publicações da literatura que abordaram, exclusivamente, o léxico e se qualificou enquanto

pesquisa com o viés cognitivista: “A variação nas terminologias pela perspectiva cognitiva: o caso da fauna e da flora” de autoria de Martins de 2022²¹; e “*La lingüística cognitiva y la lingüística aplicada a una ciencia para la comunicación transdisciplinar dos autores*” de Marulanda, Giraldo e Villegas de 2013. Apesar de constar essa característica teórica, ambos os estudos estão localizados nas tradições da Terminologia, da Linguística e dos estudos dos léxicos. O primeiro trabalho considerou que, antes da análise a partir do contexto social dos fenômenos léxicos e terminológicos, se deve buscar a compreensão dos fatores internos humanos (psicocognitivos) e suas consequências para o aparecimento e o uso das variações. No segundo artigo, os autores investigam a relação entre a unidade lexical e o seu significado em determinado contexto e durante comunicação acadêmica de uma prática médico-legal forense com a fundamentação da linguística cognitiva.

Os construtos teóricos pragmáticos estão presentes no paradigma social em CI, essa proximidade se refere ao foco na abordagem pragmática e na abordagem social, esses realçam a presença dos aspectos socioculturais e contextuais observados nos processos de comunicação do conhecimento (Weiss, Bräscher; 2017).

A **Teoria da Classificação Facetada** de Ranganathan deteve uma parcela de representação nos documentos recuperados e relativos ao aporte da construção dos tesouros. Em “A cultura e a organização do conhecimento: desafios teórico-metodológicos” de Rodrigues, em 2015, o autor, além de apresentar aspectos introdutórios, propõe uma metodologia combinada entre a Teoria de Ranganathan e a Teoria dos Níveis Integrativos – proposta pelo *Classification Research Group* – com o objetivo de mapear e estruturar aspectos teóricos do domínio da cultura. Esses aspectos para o autor são elementos estruturais tanto do conceito de cultura, quanto do seu estabelecimento, além disso, busca na organização do conhecimento os alicerces necessários para descobrir tal domínio. Nesse sentido, a Teoria da Classificação Facetada forneceu as categorias fundamentais (PMEST - Personalidade, Matéria, Energia, Espaço e Tempo) para subsidiar a identificação, os relacionamentos e a apresentação das facetas conceituais da cultura.

O mesmo aporte foi utilizado pelo autor em outro artigo de 2012, quando Rodrigues em “A Complexidade da cultura amazônica e seu reflexo para a organização e representação da informação”, busca investigar a temática da cultura amazônica através da seleção de conceitos

²¹ “A variação nas terminologias pela perspectiva cognitiva: o caso da fauna e da flora” de Martins em 2022 foi recuperado na literatura “Léxico” e “Léxico – raça, cultura, etnia”.

e de relatos orais de uma pesquisa realizada nas localidades geográficas da Amazônia, em específico nas cidades de Bragança, Castanhal e na região do arquipélago do Marajó.

Souza e Pestana (2017) apresentaram e contextualizaram as teorias dos conceitos e suas relações, dentre elas contribuições da Teoria da Classificação e as proposições rangathanianas. Dessas proposições se sobressaem o método de análise facetado. O método facetado é responsável por permitir a prospecção das faces de um assunto conforme as categorias PMEST.

Em 2005, Campos e Gomes ao elencaram os princípios da representação e da organização aplicados na construção de hipertextos, e colaboraram, significativamente, ao apontarem uma mudança na metodologia utilizada para organizar um domínio: a substituição da dicotômica/binária, desenvolvida durante séculos, para a policotomia. Segundo as autoras, a dicotômica estabelecida em dois estágios e cuja representação esquemática da Árvore de Porfírio foi deixada de lado, pois na Teoria de Ranganathan as representações dos assuntos não fazem parte de um único domínio, pelo contrário, existem complexidades que permitem a ramificação desses conforme as ramificações dos galhos de uma “Árvore Baniana” (Campos; Gomes, 2005). Em outros termos, essa teoria incorpora uma metodologia aguçada que incube um olhar complexo a favor de notar as diferentes faces de um único objeto. Nessas entrelinhas, as autoras identificaram o potencial dessa Teoria na organização de representações não lineares como os hipertextos.

Em “Terminologia e estrutura conceitual” em 2021, Gomes enfatizou na Teoria de Classificação, tida como analítica-sintética, os princípios e as categorias lógicas e ônticas de Ranganathan. A Teoria de Ranganathan e a Terminologia possuem aspectos teóricos em comuns, dentre os quais aqueles que dão visibilidade para a análise dos conceitos e, conseqüentemente, as relações com os seus termos e suas definições. Corroborando nesse sentido, Carlan e Bräscher, em 2011, colocaram a Teoria da Classificação no mesmo conjunto de teorias essenciais para estabelecer os Sistemas de Organização do Conhecimento e associá-los a interface teórica e prática da Terminologia.

O detalhamento da pesquisa de Ferreira e Maculan, em 2020, resgatou os substratos de Ranganathan e sua teoria na fundamentação teórica e metodologia dos tesauros. Esse trabalho descreveu as principais assimilações dessa teoria desde o método utilizado até a concepção do universo das ideias e da espiral dos assuntos. Essas últimas como analogias referentes a forma complexa que o assunto se apresenta para o pesquisador e usuário da informação. Foram comentados os três tipos de cânones: o plano ideacional, o plano verbal e o plano notacional. Ambos os planos objetivam perceber a gênese dos conceitos e como eles são

concebidos diante dos fenômenos da linguagem e da linguística. Outros pontos tidos como construtivos foram: a ideia de que o conhecimento é cíclico; e as categorias PMEST são o primeiro passo para inferir no modelo facetado.

A **Teoria do Conceito** é um tema de pesquisa complexo e extremamente amplo e interdisciplinar, estabelecido da relação entre muitos campos profundos e imersos em tradições históricas de longa data, das quais algumas dessas tradições não tem um consenso predefinido (Hjørland, 2009). Essa teoria conceitual da CI tem Ingetraut Dahlberg como sua fundadora e surgiu nas proposições científicas mediante as interlocuções das disciplinas filosóficas (a Ontologia, Lógica), as linguísticas e a Terminologia.

Uma das observações notadas nas argumentações de Dahlberg acerca dessa teoria recai para a importância que o conceito é analisado diante dos fenômenos das linguagens e da atividade do pensamento, de como percebemos o entorno, os objetos e os seres, e então analisá-los enquanto objetos conceituais.

Assim, é necessário compreender que os elementos dos conceitos são aferidos de acordo com o método analítico-sintético e o reconhecimento de que cada enunciado mostra atributos predicáveis do objeto, ou seja, características (Dahlberg, 1978b). Influenciada pelo método analítico-sintético, a Teoria da Classificação Facetada admite uma nova aplicação desse método por agregar os princípios da lógica, do sistema linguístico e das classificações filosóficas, bem distinta daquela defendida por Ranganathan.

Para Hjørland (2009) a melhor apreensão do conceito se deve ater às suas teorias, visualizá-las e classificá-las de acordo com as teorias epistemológicas (empirismo, racionalismo, historicismo e pragmatismo). Ou seja, cada frente epistêmica é sugerida como uma teoria do conhecimento na descoberta de seus próprios conceitos, pois alimentam as definições e o status ontológico de cada unidade conceitual dentro de oportunas possibilidades metodológicas e suas composições científicas. Stock (2010) acrescentou que o conceito e suas relações se baseiam, além do empirismo e do racionalismo, na hermenêutica, no pragmatismo e na teoria crítica.

Em paralelo dessas colocações, Hessen (2012) afirma as alterações conceituais de alguns conceitos presentes na história das escolas filosóficas e suas possibilidades do conhecimento, recordando que cada uma dessas escolas pressupõe uma investigação crítica com axiomas e conceitos fundamentais em que se exprimem a referência de nosso pensamento diante dos objetos do mundo. Muitos conceitos no decorrer da história das Ciências se manifestaram em interpretações distintas e até mesmo antagônicas.

Em muitos casos acontecem o entrelaçamento e os pontos de encontro de pensamentos que se cruzam e, ao mesmo tempo, requer um trabalho de manutenção constante das investigações científicas. A tradição da lógica formal, por exemplo, é uma delas. Apesar da necessidade da presença das tradições filosóficas, principalmente no cerne das epistemologias e dos paradigmas que modelam estudos acerca dos conceitos ambientado em campos científicos como a Linguística e Ciências Cognitiva, o inter cruzar de outras possibilidades recentes e ambientadas em epistemologias emergentes pode surtir um efeito enriquecedor e útil. Essas possibilidades são ideais para perceber assuntos como o impacto das novas representações sociais nos sistemas simbólicos e representacionais fomentados pela formulação e uso dos conceitos.

Uma das principais inquietações dessa Teoria do Conceito de Dahlberg incide na descoberta e interpretação de conceitos imbricados em conjunturas epistêmicas, e presentes em ambientes e discursos interdisciplinares, transdisciplinares e multidisciplinar. Nessa perspectiva, Hjørland (2009) questiona sobre a imprecisão da natureza dos conceitos e quais teorias são importantes no entendimento dos mesmos e aponta, inicialmente, a noção de paradigma de Thomas Kuhn como uma importante contribuição para a teoria dos conceitos.

Martines e Almeida, em 2021, teceram um diálogo com Hjørland (2009) a partir do questionamento do autor acerca das origens dos conceitos segundo a epistemologia do pragmatismo. Ambos os autores concordam com o caráter contextual, considerando os conceitos dentro dos seus respectivos domínios e, nesse ponto, que se atentam para a origem epistêmica de cada conceito e suas possíveis composições semânticas. Martines e Almeida (2021) pontuam que Hjørland é objetivo e mais assertivo no entendimento relativo à restrição significativa do conceito, pois questionou a aplicabilidade dos aspectos de veracidade que implicam na significação.

Stock (2010) afirmou que as diretrizes de construção e manutenção de sistemas de organização do conhecimento, dentre elas a ANSI/NISO Z39.19 de 2005 nos Estados Unidos e a DIN 2331 de 1980 na Alemanha, não consideram resultados da Teoria do Conceito e da Teoria das relações em toda a extensão. A partir do relato do autor, pode-se reconhecer desafios direcionados à compreensão dos conceitos e suas relações diante de questões atuais, principalmente, da Web Semântica e de possíveis formas de tratamento dessas unidades no ambiente indexado na internet.

Souza e Pestana, em 2017, introduziram a Teoria do Conceito no referencial da organização da informação em direção ao tratamento temático e descritivo, já que esses

processos demandam a modelagem de dados, do mesmo modo, são tidos em formas que agregam conceituais, entidades (autores e organizações) e assuntos (termos descritores).

Na mesma direção, Carlan e Bräscher, em 2011, e Campos e Gomes, em 2005, detalharam esse aporte durante o resgate dos aspectos teóricos dos sistemas de organização do conhecimento e esclareceram abordagens e críticas em relação a prática dessas contribuições apresentadas. Ambos os autores citaram e reconheceram o pensamento de Dahlberg, citando-a positivamente pelo menos um de seus principais artigos, dentre eles: *Teoria do conceito* (1978) e *A referent-oriented, analytical concept theory for INTERCONCEPT* (1978).

Carlan e Bräscher (2011) elencaram alguns pontos essenciais apresentados por Dahlberg: o conceito geral; o conceito individual; a intensão de um conceito (somatório de suas características); a extensão do conceito (somatória de conceitos específicos para a intensão verdadeira); a categorização e as relações entre conceitos; relação formal/categorial; relação material/paradigmática; e relações funcionais/sintagmáticas.

Gomes e Campos, em 2004, Ferreira e Maculan, em 2021, recordaram essa teoria aplicável aos tesouros. As autoras cercam suas descrições tanto nos fundamentos conceituais, quanto nas práticas de construção e manutenção esquemas entre termos. Por último, Gomes, em 2005, colocou como ponto central a Teoria do Conceito e a Terminologia, tecendo considerações de ambas as influências em pesquisas de bancos de termos e elaborações e manutenção de em tesouros, ontologias e terminologias.

Eugen Wüster foi o idealizador da **Teoria Geral da Terminologia** (TGT), destacando os primeiros princípios terminológicos a serem consolidados em normas e padrões nacionais e universalistas. Essa teoria está presente nos documentos recuperados, juntamente com a menção das duas outras teorias mencionadas anteriormente. Precisamente, as menções se referem aos estudos dos conceitos, das relações semânticas e do tratamento terminológico incorporado aos normas, padrões, diretrizes e políticas de tratamento da informação e do conhecimento. Carlan e Bräscher (2011) colaboram ao relatar a TGT como pioneira em fornecer a perspectiva prescritiva e normativa utilizada em normas e padrões relativas ao uso de termos e conceitos, sem considerar as inúmeras variações verificadas nos processos de comunicação.

O contexto de aplicação, inclusive, foi um recorte importante a ser destacado nos trabalhos analisados. Nesse tópico, pode-se notar a TGT em pesquisas teóricas e práticas focadas em: bancos de dados – Souza e Pestana em 2017; tesouro e normalização terminológica – Gomes e Campos em 2004; Terminologia e a estrutura do conceito – Gomes em 2021; sistemas de organização do conhecimento – Carlan e Bräscher em 2011;

Souza e Pestana (2017) encontram limitações nas contribuições de Wüster e cuja base está nessa teoria, se referindo: à hipótese de que o conceito preexiste à expressão; e à universalização do conhecimento técnico e científico ao considerar a estruturação uniforme de uma área como única em todos os grupos e contextos. Nesse sentido, a Teoria Geral é normativa e concebe o termo de modo unívoco por excluir fatos e variáveis culturais e históricas presentes nas comunicações em especialidades.

Em Gomes e Campos, 2004, e Gomes, 2021, pode-se reconhecer o método onomasiológico como sendo um método particular, introduzido nessa teoria e interiorizado como um dos principais aportes metodológicos linguísticos e terminológicos presentes na construção de tesouros e vocabulários controlados. A abordagem onomasiológica tem origem na onomasiologia oriundos dos estudos da Semântica Linguística.

A abordagem e o método onomasiológicos se iniciam com base no contexto de um conceito e procuram, dentre os seus significados, aquele que recebe tal designação, diferentemente da abordagem e do método semasiológico que elege uma palavra ou um termo na qualidade de item a ser analisado, considerando essa unidade polissêmica e cujo sentido se manifesta no contexto (Gomes, 2021). Assim, percebe-se o método onomasiológico como um processo prático e influente nas etapas de seleção e de registro das unidades terminológicas, perpassando pela sua função de identificação de significados para, posteriormente, estruturar as relações semânticas em seus contextos e esquemas nocionais.

Gomes e Campos (2004) enfatizaram que esse método ainda reconhece o conceito como unidade do tesouro materializada no termo, destacando que a Teoria da Terminologia postula que os termos devem ser monorreferenciais. Nesse raciocínio, não podemos considerar a palavra enquanto unidade no tesouro, pois a palavra são elementos plurireferenciais e polissêmicos e requer a tradução ou o controle terminológico para corresponder ao conceito e à forma padrão, livre de dispersões lexicais.

Em “*Empiricism, rationalism and positivism in library and information Science*” de 2005, Hjørland analisou e descreveu a influência epistemológica do empirismo, racionalismo e positivismo presentes na Ciência da Informação e Biblioteconomia. A apresentação desse estudo emitiu um anseio de verificar os horizontes e as possibilidades viáveis de aplicação e reorganização do campo de pesquisa e de atuação frente à superação de paradigmas já consolidados e à preocupação constante em revisar os discursos científicos. Amorim, Viana e Bräscher, em 2019, igualmente, percorreram o mesmo caminho ao elencarem os aspectos epistemológicos em estudos sobre análise de domínio sob as lentes do racionalismo e do empirismo.

Os dois estudos abrem uma discussão densa e esboçada nos estudos recentes da CI, e mais precisamente, nos alicerces da proposta da Análise de Domínio. Por outro lado, Hjørland em 2009 e Stock em 2010 trataram de temas específicos nas teorias dos conceitos e seus desdobramentos. Nesse limiar, Hjørland, em 2022, apresentou a segunda parte de uma trilogia sobre ciência, detendo-se nos vários campos de estudo dos estudos científicos e recordando que o racionalismo, o empirismo, o historicismo e o pragmatismo estão sempre presentes em todas as disciplinas paradigmas.

Primeiramente, concebe-se que o estatuto das tradições filosóficas e suas caracterizações epistemológicas sugerem uma linha do tempo formada por sobreposições, oposições e articulações históricas marcantes na Filosofia, na Sociologia da Ciência, na Sociologia do Conhecimento e nas Teorias do Conhecimento. Na CI, não é diferente, tendo em vista dois fatos marcantes: a herança das tradições das ciências exatas e biológicas e da busca pelo rigor científico nas ciências sociais e humanas. Duas escolas tradicionais se destacaram a priori nesse panorama: o racionalismo e o empirismo.

O **racionalismo** tem início com René Descartes (1596-1650) após um contexto marcado por escolas anteriores que preconizavam à realidade, às coisas abstratas e concretas e Deus. Descartes por outro lado, entendia que a prioridade deveria incidir no humano e no exercício do pensar como premissa basilar nas investigações filosóficas. Esse filósofo e, também matemático, influenciado pelas leis da Matemática, em especial a Geometria, afirmava que o caminho científico a ser percorrido deveria iniciar no ato do raciocínio e somente através da razão o método poderia confrontar e desnudar toda e qualquer inquietação humana e científica. Além de Descartes, Benedict de Spinoza (1632-1677) e Gottfried Wilhelm Leibniz (1646-1716) foram os maiores expoentes do racionalismo (Hjørland, 2005).

O método defendido pelo racionalismo é destinado à redução de qualquer problema aquilo que não pode ser questionado (afirmações em evidência), das quais as afirmações evidentes devem ser passíveis de comparações e de intervenções para admitir que novos conhecimentos podem ser deduzidos (Hjørland, 2005).

Analisando essas proposições racionalistas, é plausível questionar quais elementos conceituais e teóricos viabilizam o julgamento e a verificação de um conhecimento. Para Hessen (2012) o conhecimento no racionalismo só tem tal denominação de conhecimento se for possível atribuí-lo enquanto universal e autêntico. No caso da CI, temos um exemplo desse exercício epistêmico apontado por Hjørland (2009) ao declarar que a análise do racionalismo tende a usar a proposta “de cima para abaixo” no processamento da informação, em outras

palavras, abordar um determinado conjunto de dados, a priori, a partir das categorias preestabelecidas.

O **empirismo**²² foi outra tradição epistemológica que incorporou uma nova compreensão do conhecimento no estatuto científico da ciência moderna, principalmente, as ciências biológicas e experimentais. Fundado no âmbito britânico, o empirismo clássico tem como seus principais representantes John Locke (1632-1704), George Berkeley (1685-1753), David Hume (1711-1776) e John Stuart Mill (1806-1873), e, já no século XX, Bertrand Russell e os empiristas lógicos (Hjørland, 2005).

O surgimento do empirismo atestou a incorporação de uma noção para além do ato racional, ampliando a busca pelo conhecimento a partir da relação do homem com os objetos e a vivência perceptiva consequente desse relacionamento. A tese empirista postula que a única fonte do conhecimento humano é aquela adquirida por meio da experiência, sendo a antítese do pensamento do racionalismo. É nessa premissa que Hessen (2012) esclarece que a cada pessoa por ocasião do nascimento está vazio de conteúdos e se configura como uma tábula rasa, uma folha em branco, onde a experiência irá escrever e todas os relatos e conceitos serão captados, até mesmo os mais universais e abstratos. Hjørland se aproxima de Hessen e acrescenta essa *modus operandi* na aquisição e constituição dos conceitos:

O desenvolvimento consistente desta visão durante a história da filosofia levou a um número abrangente de suposições sobre o mundo (ontologia), sobre a mente (psicologia), sobre conceitos e linguagens etc., que formam a base do ponto de vista empirista. Por exemplo, **a visão de que um ser humano recém-nascido não tem conhecimento ("tábula rasa"), que conceitos complexos são criados na mente pelas combinações de conceitos mais simples correspondentes a sensações simples e que os processos de aprendizagem são governados por leis de associação etc.** (Hjørland, 2005, p. 134, tradução nossa, grifo nosso).

Desse modo, a ciência utiliza o experimento como principal meio e capacidade de constar e formular seus construtos e obter sua validade. Em consequência, os conceitos seriam resultantes do processo que se inicia na experiência e nas sensações para alimentar os processos cognitivos interiorizados em cada sujeito. Hjørland (2005, 2009) descreveu as epistemologias empiristas se consolidaram segundo as lentes de um observador, cujo processo científico é concebido como coleta de observações a serem apuradas e o método é o da indução. Na prática o empirismo e o racionalismo são notados, da mesma forma, na oposição dos métodos: indutivo (empirista) *versus* dedutivo (racionalista).

²² Hjørland (2005) é claro ao diferenciar estudo empírico e empirismo, já que ambos os conceitos detêm definições particulares, tratando o empirismo como uma tradição e vertente das teorias do conhecimento.

Na ocasião do entendimento desses métodos, Amorim, Viana e Bräscher, em 2018, mapearam 50 pesquisas da Análise de Domínio pertencentes a literatura da CI, identificando que 29 de matriz empirista e 13 de matriz racionalista. Os autores inferiram algumas considerações.

[...] destacam-se **os métodos utilizados, sendo que particularmente chama a atenção a associação dos estudos métricos com a matriz empirista**, quando tradicionalmente nas ciências, tais métodos estão associados ao racionalismo, o que merece maior aprofundamento. Em geral, o racionalismo se baseia na matemática e o empirismo nas ciências naturais e na observação direta de experiências repetidas, além do fato de que nesta vertente, a indução substitui a dedução (Amorim; Viana; Bräscher, 2018, p. 45, grifo nosso)

No limiar da caracterização metodológica, as duas tradições se mostraram atuantes e influentes para diferentes frentes de pesquisa. É o caso dos estudos métricos que apesar de pressupor relações diretas e estreitas com as ciências exatas – dessas em que a herança racionalista é influente, como na Estatística, na Matemática e na Ciência da Computação destinadas ao tratamento dos dados e das linguagens – podem agregar a matriz empirista no aproveitamento de seus métodos.

Nesse interim, os estudos métricos, por terem se consolidado e fundamentados no século XX, apresentaram uma configuração segundo métodos e heranças epistemológicas a partir das formalidades de perspectivas que incubem analisar um cenário científico de discursos interdisciplinares. Isso ocorre onde os métodos de influência empirista se sobressaíram daqueles racionalistas, por permitirem a experimentação das leis métricas e uma adaptabilidade em relação as tecnologias de comunicação e meios informacionais dinâmicos, dos quais perpassam os indícios quantitativos e qualitativos (dados científicos e bibliográficos).

Quanto à distribuição de métodos utilizados nas pesquisas de Análise de Domínio em CI caracterizado pela matriz racionalista, Amorim, Viana e Bräscher (2018) destacaram a categorização, a modelização (aplicação de um grupo de classes, categorias, modelos ou facetas em um movimento dedutivo) e a mineração textual.

Podemos conferir na citação abaixo o empirismo no cenário dos estudos de usuários:

Na psicologia (e nos estudos do usuário na Ciência da Informação), uma forma dominante de empirismo / positivismo tem sido o behaviorismo. Esta é uma visão que implica que os usuários estão respondendo aos estímulos de maneiras mecânicas, seguindo leis universais que são comuns a todos os seres humanos (ou mesmo para animais superiores ou para todos os animais). **O empirismo e o behaviorismo são, portanto, visões que tendem a negligenciar o papel da cultura e da linguagem nos processos cognitivos (uma visão que compartilham com o racionalismo e com a**

maioria das chamadas ciências cognitivas). A implicação para estudos empíricos é, obviamente, que um estudo de uma população de usuários é visto como representativo dos usuários em geral (e que estudos sistemáticos de diferenças entre diferentes grupos de usuários são relativamente negligenciados) (Hjørland, 2005, p. 140, tradução nossa, grifo nosso).

Na análise dos métodos introduzidos na elaboração de linguagens documentárias, mais recentemente os sistemas de organização do conhecimento, notou-se que o empirismo e o racionalismo estiveram presentes nos princípios e ordenação de esquemas classificatórios e relacionamentos entre os assuntos desde as abordagens *top-down* (dedutiva) e *bottom-up* (indutiva), sendo encaminhadas em direção ao exercício da lógica e da razão até as concepções universalistas que inspiraram as classificações universais e decimal (CDU, CDD).

Mostafa e Cruz 2010 enfatizaram o quanto é comum as linguagens (os tesouros e outras especializadas) compor suas relações hierárquicas a partir da lógica aristotélica – uso da dedução lógica, dos silogismos e da análise dos predicados como gênero, espécie e diferença. Assim, nessas questões, o Empirismo inglês de David Hume se sobressai enquanto uma referência adequada às relações associativas em linguagem documentária. As autoras exemplificaram de modo prático como a visão empirista de Hume pode auxiliar no tratamento de termos e suas relações associativas. Essa contribuição do empirismo de Hume também esteve como principal vertente descrita por Amorim, Viana e Bräscher, em 2019, cujas colocações esboçadas podem inferir que os sentidos e o contato direto com o cotidiano e os fatos são instrumentos metodológicos imprescindíveis na CI.

O **positivismo** alcançou uma amplitude nos estudos das ciências sociais do século XIX e XX. Sua origem recebeu traços ideológicos e formulações do filósofo francês August Comte (1798-1857), seguindo caminhos diferentes na Inglaterra e no positivismo lógico da Áustria e da Alemanha (Hjørland, 2005). Sua centralidade ainda preservou o humano e sua percepção como o elemento primeiro da Filosofia e a procura por constatar verdades universais que garantissem a estabilidade e as conformidades que a ciência buscava instaurar: o sujeito pesquisador protagonista de uma realidade que ele mesmo assegura desvendar de acordo com os métodos e os mecanismos institucionais (o discurso, as disciplinas etc.). Complementando, esse ponto de vista, foi definido que a ciência deveria se ater ao que é dado de forma positiva e aos fatos perceptíveis através da experiência e não considerar a existência da sabedoria individualizada ou o conhecimento filosófico-metafísico, mas somente aquele conhecimento provindo das ciências particulares (Hessen, 2012).

O positivismo não esteve presente somente nas ciências sociais, já que uma de suas máximas eram imprimir um estatuto de ciência universal, dos quais situavam todas as ciências

no mesmo patamar. Outro tópico relevante está em tratar o conhecimento através da linguagem e do discurso, seguindo regras claras e critérios de verificação e mensuração em uma realidade objetiva e que está para além da mente humana. Ou seja, em uma dimensão empírica e na dinâmica do mundo.

Com Hjørland, em 2005 e 2022, é possível notar uma expressiva presença da vertente do positivismo lógico na CI. Após uma série de contestações que fomentaram essa nova proposta científica, possibilitou-se um olhar lógico diante de objetos e fenômenos sociais, dos quais poderiam ser reduzidos e traduzidos em variáveis e em unidades analíticas. Nesse trecho abaixo, compreendemos melhor o ponto de partida dessa vertente e suas respostas para as concepções empiristas e racionalistas:

Os positivistas lógicos argumentaram que o conhecimento sensorial é o tipo mais certo de conhecimento e, portanto, qualquer conceito não diretamente relacionado à experiência sensorial deveria ser "traduzível" em conceitos observacionais. Esses conceitos que não podem ser traduzidos são vistos como sem sentido. Essa afirmação levou a uma visão dualista da ciência: uma parte observacional e uma parte teórica. Eles combinaram uma visão racionalista e uma empirista, exigindo uma separação distinta entre os componentes formais e empíricos da ciência. **A oposição metafísica anterior de racionalismo e empirismo foi transformada em uma distinção linguística entre verdades analíticas e sintéticas** (Hjørland, 2005, p. 138, tradução nossa, grifo nosso).

Conforme o relato, Hjørland resumizou um dos principais pontos do positivismo lógico aquele que priorizava os objetos e os fenômenos científicos através de unidades conceituais e elementos discursivos. A prática de indexação, por exemplo, analisada em comparação as prerrogativas positivistas, pode implicar em alguns processos e princípios alinhados às observações lógicas de pesquisadores e indexadores preocupados em aplicar somente técnicas e métodos analíticos. O entusiasmo maior estaria, então, mais próximo de atestar a aplicação de conhecimentos específicos – esses estabelecidos do diálogo com a Linguística e a Filosofia – no tratamento de unidades semânticas e somente, assim, garantir soluções práticas de controle, disseminação e organização das coleções mantidas, quase que, com exclusividade por instituições científicas e governamentais.

Hjørland (2022) assinalou que Karl Popper (1902-1994) e Thomas Kuhn (1922-1996) – duas figuras da ciência do século XX – criticaram as afirmações do empirismo e do positivismo, dos quais Popper, mesmo associado ao positivismo lógico, negou o inducionismo e a crença que preconizam a acumulação das observações e suas inferências pautadas em princípios universais; e Kuhn, visto como um autor de interpretação hermenêutica, negava o processo científico pela acumulação de fatos, já que para ele os fatores poderiam ou não serem

articulados em uma trajetória evolutiva, sendo possíveis de receber a denominação de “paradigmas”.

Os conflitos entre empirismo e racionalismo parecem no positivismo terem uma nova perspectiva de mudança e superação. Não que a versão positivista tenha sanado por completo as questões ainda estáveis nas ciências, mas, interpretando Hjørland (2005; 2022) notou-se um desejo expresso dos positivistas em desvendar o conhecimento conforme seus aspectos linguísticos e lógicos.

As críticas e as interpretações do positivismo lógico foram elencadas por Hjørland (2005) e se concentraram na citação de dois autores em particular: Svenonius (2004) ao explicar o operacionalismo como uma teoria do significado que define um conceito, em muitos casos, enquanto uma variável; Kincaid (1998) ao apontar que as pesquisas sociais utilizam métodos estatísticos para constatar hipóteses e, de forma implícita, elaborar inferências estatísticas em um processo mecânico e puramente lógico, além de salientar as correlações entre variáveis sem implicar em conclusões acerca de suas causas.

Freitas, Bufrem e Breda (2016) elencaram opções metodológicas em pesquisas da CI que contribuíram com a aplicabilidade da Análise de Domínio, dos quais foram observados a ocorrência dessas opções e a elucidação das correntes epistemológicas – como os estudos afins ao positivismo na década de 1970 e, na década de 1980, a visualização de uma influência da Fenomenologia que ampliou as opções metodológicas nas pesquisas qualitativas. Apesar de confundirem o empirismo, o positivismo e o realismo, estas escolas são por natureza fortemente antirrealistas, em detrimento disso deve ser feita uma distinção nítida entre empirismo e realismo (Hjørland, 2004b).

Thellefsen (2023), ao investigar os níveis que formulam o conceito de informação, discutiu e esclareceu a base realista da Análise de Domínio e da Organização do Conhecimento, declarando em suas considerações que o realismo crítico pode conciliar a disputa entre concepções externalistas e internalistas do conceito de informação, de modo que se aproxima do realismo filosófico sugerido por Hjørland.

A premissa do **realismo epistemológico** é de que existe uma realidade independente da mente e o realismo não deve ser confundido com a tese de que todas as afirmações científicas são verdadeiras ou a ocorrência da ingenuidade perante as afirmações de conhecimento (Hjørland, 2004b). Essa noção compreende que os objetos e entidades têm sua existência sem a percepção ou a interferência do pensamento humano. Ou seja, os seres vivos e as espécies vivem sem que uma teoria ou classificação taxonômica possam denominá-los e distingui-los uns dos outros.

Em oposição a essa noção de realismo encontramos um pensamento antagônico caracterizado por **antirrealismo**. Hjørland (2004b) comentou que existe uma presença generalizada do antirrealismo na CI e na Biblioteconomia em muitas pesquisas de recuperação da informação e estudos baseados na abordagem cognitiva.

Outras perspectivas epistemológicas e teorias do conhecimento presentes na literatura foram esclarecidas por Hjørland (2005):

- a) O **historicismo** baseia-se em desenvolvimentos históricos e na compreensão de observações, sejam aquelas que envolvem as teorias, fenômenos ou objetos sociais, além de ser influenciada por fatores e questões culturais;
- b) A **hermenêutica** pertence ao Historicismo, antes não era considerada uma teoria do conhecimento, mas uma teoria de interpretação de textos. Mantém uma relação estreita com as humanidades, pois entende que a pré-compreensão humana é imprescindível para interpretação dos discursos e quaisquer objetos textuais;
- c) O **pragmatismo** está baseado na suposição de que o conhecimento não pode ser neutro, sendo importante descobrir os valores e as consequências inerentes a qualquer afirmação de um conhecimento. Tem afinidade com o historicismo, por, também, compreender que as observações se caracterizam de acordo com contexto e estão “carregadas de teoria”, mas diferem ao tentar adquirir um posicionamento explícito relativo ao propósito da pesquisa e do processo cognitivo do qual envolve.

Essas explicações de Hjørland foram citadas por Stock em 2010 para elucidar tópicos levantados acerca da noção de conceito e suas relações semânticas. Conforme a leitura de Stock, compreendeu-se que essas relações semânticas são baseadas nas teorias e noções epistemológicas, pois, como verificamos na explicação de Hjørland cada uma delas ensina por métodos e formas distintas de enxergar um determinado conhecimento. Exemplificando, retoma-se o pragmatismo – indicado para entender um conhecimento social enraizado em tradições que não são filosóficas ou das ciências exatas – no caso dos estudos dos léxicos ambientados em países com forte presença de imigrantes estrangeiros. Nesse exemplo, a percepção tende a desvendar materiais de pesquisa com ênfase em pesquisa que utilizam falantes, discussões históricas, culturais e de natureza teórica da linguística para possibilitar a elaboração de políticas linguísticas ou materiais linguísticos.

As argumentações apresentadas por Hjørland, em 2004b, se mostraram como um presságio ou uma construção dialógica para introduzir o que o autor chama de paradigma da Análise de Domínio, considerando que ele busca discutir amplamente as bases e escolas filosóficas presentes na CI. Inicialmente, o autor faz uma crítica em relação a influência predominante do antirrealismo nos discursos científicos e aqueles que ainda questionam a existência de uma realidade independente da mente. É nítido a cristalização de algumas hipóteses formalizadas em questões pragmáticas, dando uma ideia de que a epistemologia da CI deve ser constantemente avaliada, mesmo que diante de teorias já consolidadas.

Nesse sentido, Hjørland em 2004b e em 2005, enfatizou a necessidade de uma análise crítica e de uma revisão dos suportes teóricos e das tradições epistêmicas que tendem a modificar abertamente as abordagens, considerando a natureza complexa, evolutiva, Inter e multidisciplinar dos fenômenos e objetos de estudo da CI e da Biblioteconomia. Podemos visualizar detalhadamente a descrição de Hjørland acerca das influências das tradições epistemológicas e filosóficas no entendimento das teorias do conceito.

Quadro 9 – As Teorias do Conceitos e as Teorias do Conhecimento em Hjørland

Racionalismo	O racionalismo é o ideal de basear o conhecimento em lógicas, princípios, regras e modelos idealizados. O racionalismo é cético em relação ao empirismo e às experiências sensoriais que não são organizadas de acordo com princípios, que são, de uma forma ou de outra, a priori em relação à experiência (por exemplo, ligados rapidamente aos nossos sistemas cognitivos). O racionalismo pode parecer estranho de um ponto de vista empírico moderno. É, no entanto, uma forte posição clássica com argumentos que não podem ser ignorados: cada processo de obtenção de informação é sempre uma interação de processamento ascendente e descendente, onde o racionalismo é necessário para explicar este último tipo de processamento de informação. A realidade é interpretada em conceitos e categorias lógicas.⁴³ A “divisão lógica” é um método racionalista (e na organização do conhecimento a abordagem analítico-facetada é a que mais se aproxima deste ideal). A investigação empírica inspirada em modelos teóricos deve ser considerada racionalista se estes modelos não forem eles próprios derivados empiricamente.⁴⁴ O racionalismo entende os conceitos como anteriores à experiência sensorial e como formadores de estruturas lógicas. Conceitos complexos são analisados, decompostos e definidos de acordo com os conceitos simples que os constituem (Hjørland, 2009, p. 1524, tradução nossa, grifo nosso).
Empirismo	O empirismo é o ideal de basear o conhecimento em observações (e em induções de um conjunto de observações). Devem ser evitadas seleções teóricas e interpretações de observações. (Se não, passamos para outra abordagem). As observações são vistas como “dadas”, não como contextuais ou dependentes de teoria. As abordagens empiristas baseiam-se na análise ascendente (a informação é lida a partir dos objetos, a percepção é vista como um processo receptor ou passivo). O empirismo clássico criticou a doutrina aristotélica dos conceitos como características essenciais abstratas das coisas.³⁸ O empirismo argumenta que os conceitos são formados na aprendizagem com base em relações de semelhança entre as coisas e através de convenções adquiridas ou aprendidas entre coisas e palavras (Hjørland, 2009, p. 1523, tradução nossa, grifo nosso).

Positivismo lógico	O positivismo lógico (e o empirismo lógico) pode ser considerado uma combinação de empirismo e racionalismo caracterizada por reduzir conceitos a combinações de atributos sensoriais e, assim, ignorar as dimensões histórico-culturais e pragmáticas dos conceitos . Foi influente na primeira metade do século XX, mas perdeu influência por volta de 1962, quando Kuhn publicou Estrutura. Não consideraremos mais esta visão porque os princípios mais importantes subjacentes esta visão está presente em forma pura no empirismo e no racionalismo (Hjørland, 2009, p. 1525, tradução nossa, grifo nosso).
Pragmática	O pragmatismo entende os conceitos como uma forma de fixar partes da realidade no pensamento, na linguagem e em outros sistemas simbólicos. Estas partes da realidade não são fixadas apenas pela semelhança (como suposta pelo empirismo), pela divisão lógica (ou regras semelhantes como assumidas pelo racionalismo), ou pela genealogia (como suposta pelo historicismo), mas pelo que é considerado como classes equivalentes funcionais das coisas. O conceito “comida” é um bom exemplo porque o que é considerado comida não se baseia na semelhança, mas na função. Tipos muito diferentes de coisas podem servir como alimento, e o conceito de alimento deve ser entendido como criado para as pessoas lidarem com o mundo. Todos os outros conceitos também são vistos de forma semelhante (incluindo o conceito de “conceito”). A visão pragmática dos conceitos foi formulada pela primeira vez por William James. Uma definição mais recente de conceitos da perspectiva pragmática foi formulada pelo cientista da computação John F. Sowa (Hjørland, 2009, p. 1526, tradução nossa, grifo nosso).
Historicismo	O historicismo difere tanto do empirismo quanto do racionalismo em aspectos importantes. Nas palavras de Mammen (2008, p. 25), “escapou da armadilha do empirismo-racionalismo”, segundo a qual o nosso conhecimento individual deve vir dos sentidos ou ser inato. Os conceitos devem, portanto, evoluir a partir do processamento de informações sensoriais de acordo com mecanismos cognitivos. Esses mecanismos são, pelo empirismo e pelo racionalismo, considerados biologicamente dados, e não desenvolvidos cultural ou socialmente. A fuga da armadilha do empirismo-racionalismo é, portanto, a compreensão de que as experiências relacionadas com a aprendizagem de línguas e outros sistemas simbólicos mudam o nosso sistema cognitivo de maneiras fundamentais. ⁵¹ De acordo com teorias historicistas, como a teoria da atividade, a aprendizagem de um conceito (digamos [uma xícara]) não é apenas associar a palavra “xícara” (som ou imagem) com objetos com algumas combinações de propriedades físicas e químicas (como supõe o empirismo). Além disso, não combinamos um input sensorial com um conjunto de categorias lógicas (como supõe o racionalismo). Aprendemos conceitos com “profundidade histórica”, ou seja, aprendemos algo sobre a função das xícaras e sobre as experiências acumuladas da humanidade em relação às xícaras e suas funções. Isto também significa que as concepções são culturalmente relativas e só podem ser compreendidas num contexto cultural (Hjørland, 2009, p. 1526, tradução nossa, grifo nosso).

Fonte: Hjørland (2009).

Essas teorias sugerem diferentes formas e pontos de vista. Elas são imprescindíveis para os exercícios ontológicos que buscam verificar a essência e a substância de algo, no caso da CI possuem um valor analítico na compreensão de unidades conceituais (termos, conceitos ainda em fase de elaboração, conceitos e definições de domínios, axiomas, expressões sintagmáticas, anéis de sinônimos, modelos conceituais e outros).

A literatura em questão expressou relações entre as disciplinas mencionadas e indicou as contribuições da **Semiologia** ao reintroduzir as teorias do signo em consonância com

as investigações do binômio: conceito-termo. Verificou-se esse destaque, principalmente, na aplicação de alguns fundamentos já consolidados nos tesouros, como a relação entre os conceitos e os objetos que uma disciplina ou assunto temático buscam formaliza-los através de termos descritores. Essa também pode ser concebida em outras publicações da AD com questionamentos semelhantes àqueles da constituição e atualização das linguagens de especialidades.

Martines e Almeida (2021) afirmaram que a interação com a Semiologia é uma via direcionada à compreensão da significação e das funções de cada elemento participante de tal processo interpretativos e de atribuição de sentidos. Aqui a principal vertente identificada com o relacionamento mais estreito, em termos interdisciplinares, é a das Linguagens Documentárias, apesar de que há indícios da presença dos fundamentos semiológicos nos trabalhos publicados na disciplina da Organização do Conhecimento e outras da CI. É nesse sentido que Saldanha e Souza (2017) afirmaram que a pragmática e “semiótica do pensamento dos tesouros” se mostraram como ferramenta teórica localizada no centro das discussões contemporâneas da epistemologia da Organização do Conhecimento, principalmente nos planos social e cultural.

Nesse sentido em Martines e Almeida, foi entendível o papel da Semiologia e da **Semiótica de Peirce** para as noções de signos evidenciadas na fundamentação teórica e prática das Linguagens Documentárias, dentre elas os tesouros:

Para a semiologia, o signo assume um valor semântico-linguístico produzido pelo estrato social e a compreensão de sua significação é decorrente da análise da imanência desses valores. Para a semiótica peirceana, o signo assume um papel lógico-cognitivo, pois sua sistematização entre signo, objeto e interpretante promove a semiose. Somado a isso, ainda é possível apresentar uma terceira concepção sobre o signo, uma perspectiva ideológica (Martines, Almeida, 2021, p. 700).

Apesar de ambas terem como objeto o signo, a **Semiologia** descrita é aquela articulada pelo filósofo e linguista suíço **Ferdinand de Saussure** (1857-1913) em sua busca de interpretar os signos e suas diversas manifestações na sociedade, sejam elas faladas ou articuladas em imagens e outras expressões, e cuja função está em “[...] compreender o valor semântico proposto pelo signo” (Martines, Almeida, 2021, p. 700). Saussure estabeleceu um aporte teórico estruturalista inspirado nos seus próprios princípios constituintes da clássica e histórica Linguística Estruturalista.

A Semiótica Pierciana, por outro, lado se refere à teoria dos signos defendida pelo filósofo, e linguista, estadunidense Charles Sanders Peirce (1839-1914) que situou suas premissas na lógica formal e nas observações acerca das representações dos objetos. E por

último, a terceira concepção é descrita como aquela que considera a ideologia de um signo conforme a **Semiologia do filósofo russo, Mikhail Mikhailovich Bakhtin**.

Diante dessas três contribuições, Martines, Moreira e Almeida (2022, p. 53) revelaram que os recursos complexos e eficientes da Semiótica Peirceana possibilitam “ [...] mensurar o impacto do signo, atuando como conceito e como tesouro, já que além de estar alinhado às categorias fenomenológicas, também é avaliado pelo pragmatismo, o que efetiva o conceito e o tesouro como um hábito, tendo condições de direcionar uma conduta para agir no futuro”.

Marcondes (2021) inferiu que a Ciência da Informação e a Organização do Conhecimento se aproximam e se fundamentam em teorias como a Ontologia – Filosofia – e a Semiótica de Peirce. O autor verificou uma proximidade entre a tríade da teoria dos signos na visão de Peirce (*representamen*, objeto e interpretante) com o triângulo do conceito da representação formal de uma unidade de conhecimento de Dahlberg (1978a; 1992) composto pelo referente (A), as características (B) e a forma verbal (C). Nessa analogia, Marcondes entende outros pesquisadores da CI e Organização do conhecimento no uso desses dois esquemas triádico:

Os autores **percebem várias similaridades entre a Semiótica Peirceana e a Teoria do Conceito de Dahlberg, como estarem inseridas nas teorias da linguagem, usarem o triângulo para relacionar seus elementos e suas interações, incorporarem tipologias e subclassificações destes elementos e endereçarem a questão da representação do conhecimento**. Enquanto em Peirce a percepção/cognição é um aspecto importante da sua teoria e um fundamento para se pensar na representação do conhecimento, este aspecto não é endereçado na Teoria do Conceito; nesta os aspectos comunicacionais e consensuais são destacados. Na teoria de Peirce o significado é o resultado final do processo de semiose, o aspecto consensual de um conceito, que Peirce chama de interpretante final. Na teoria da Dahlberg este aspecto consensual, o significado do conceito, é dado pelos diferentes tipos de relações entre conceitos, obtidas das proposições verdadeiras feitas sobre o objeto e que sugerem estas relações entre conceitos. Segundo os autores, estas relações permitem, na Teoria do Conceito, classificar e relacionar conceitos. (Marcondes, 2021, p. 272, grifo nosso)

As nuances percebidas nessas contribuições, embora avaliadas como duas teorias com propósitos particulares, podem ser identificadas na prática da descrição física e temática. No estudo dos aspectos semióticos na criação de registros bibliográficos, Reis, Fujita, Santos e Zafalon, em 2018, acrescentaram mais duas vertentes teóricas dos signos dentre as cinco principais:

Charles Sanders Peirce (1839-1914) entende que o signo é triádico, composto por um Interpretante, um Representamen e um Objeto. Já na visão de Ferdinand de Saussure (1857-1913) o signo é diádico, composto por Significante e Significado. Para Charles

W. Morris (1901-1979) existe o Signo triádico proposto – em termos de Sintática, Semântica e Pragmática. Na visão de Roland Barthes (1915-1980) o “signo” diádico pode ser entendido em termos de denotação e conotação. **Thomas A. Sebeok (1920-1991) realizou o estudo de semiótica para incluir significação e comunicação de sistemas não humanos. E finalmente, Umberto Eco (1932-2016) propôs o termo signo como “produção de signos”.** (Friedman 2012, p. 131 apud Reis et. al, 2018, p. 46, grifo nosso)

Após os teóricos clássicos, duas outras frentes semióticas detiveram atenção: a percepção de Thomas Sebeok e aquela que retoma o universo dos signos e sua dimensão não humana – visto que a comunicação e as interações de significados se confrontam com outras dimensões, além da esfera humana.

A publicação “Tesouro do CDI-SEBRAE/ES” de Corrêa e Thomes, em 2022, foi qualificada como um estudo descritivo que utilizou os princípios da Semiótica para guiar os procedimentos de construção do tesouro do Sistema Sebrae (Serviço Brasileiro de Apoio às Micro e Pequenas Empresas) do estado brasileiro do Espírito Santos. Esse estudo teve como um dos objetivos a aplicação da Semiótica na verificação das relações componentes semânticos.

As publicações “*Competencia léxica y procesos de variación semántica en los verbos del parlache*” de Guerrero, Montoya em 2022 e “*Reflexiones en torno a la situación sociolingüística de las lenguas criollas de base léxica inglesa del caribe*” de León, de 2014, foram os poucos trabalhos recuperados na literatura dos léxicos que abordaram a **Sociolinguística**. O primeiro estudo de Guerrero, Montoya, em 2022, reiterou a sociolinguística como aporte teórico e metodológico para averiguar como se dá a competência léxica dos usuários de uma população latino-americana através de variações identificadas em verbos e suas atribuições de significado. Daí a exploração metodológica esteve fixada no léxico e no âmbito sociocultural de seus falantes, demonstrando a influência dessa disciplina no campo dos estudos lexicais.

Para Moraes (2018) a toponímia indígena tupi é tida como a marca registrada da historiografia indígena, pois está presente em diversos estados e cidades brasileiras, sendo um exemplo do interesse de estudo da Sociolinguística.

A importância do estudo de León, em 2014, se concentrou nas reflexões que buscaram descobrir as origens e as atuais características das línguas criolas de base inglês do Caribe. O aporte linguístico para entender o contexto cultural e étnico-racial dessas línguas utilizou uma análise social e buscou por elementos históricos. Análise e elementos que Sociolinguística elegem como essenciais.

Contudo, nesses dois últimos artigos não foram inferidas as relações da Sociolinguística com tesouros ou detalhes que pudessem implicar, com exclusividade, na

utilização em pesquisas acerca dos léxicos indígenas. Mostrando, assim, poucos indícios e discussões esclarecedoras das técnicas ou abordagens teóricas e metodológicas da Sociolinguística.

6.3 Aspectos metodológicos

A concentração de publicações acerca dos aspectos metodológicos tem sua maior parcela em estudos empíricos para desenvolvimento de tesouros. Os demais artigos tratam de metodologias em estudos de Análise de Domínio, Lexicologia, Terminologia e estudos críticos e empíricos da CI, salientando também pontos convergentes e divergentes daqueles apresentados na construções e diretrizes de vocabulários estruturados. Conforme a terminologia utilizada pelos autores, identificou-se abordagens, métodos, ferramentas, padrões, modelos, propostas metodológicas descritas nos quadros desta seção.

Quadro 10 – Abordagens e análises

Aspectos metodológicos	Análise de Domínio	Tesouros	Etnia, Raça, Cultura e Léxico	Léxico	Etnia, Raça, Cultura e Análise de domínio	Tesouros e Análise de Domínio	Etnia, Raça, Cultura e Tesouros
Abordagem centrada no usuário	Mai, Jens- Erik - 2005 Martín-Moncunill, David; Sicilia-Urban, Miguel-Ángel; García-Barriocanal, Elena; Sánchez-Alonso, Salvador – 2015	Shiri, Ali Asghar; Revie, Crawford; Chowdhury, Gobinda - 2002 Mai, Jens- Erik – 2005; Yoon, JungWon – 2009					
Abordagem orientada ao documento	Mai, Jens- Erik – 2005	Mai, Jens- Erik – 2005					
Abordagem centrada em documento	Mai, Jens- Erik – 2005	Mai, Jens- Erik – 2005					
Abordagem centrada em domínio	Mai, Jens- Erik – 2005; Martín-Moncunill, David; Sicilia-Urban, Miguel-Ángel; García-Barriocanal, Elena; Sánchez-Alonso, Salvador – 2015;	Mai, Jens- Erik – 2005					

Aspectos metodológicos	Análise de Domínio	Tesauros	Etnia, Raça, Cultura e Léxico	Léxico	Etnia, Raça, Cultura e Análise de domínio	Tesauros e Análise de Domínio	Etnia, Raça, Cultura e Tesauros
Abordagem da epistemográfica	Moraes, Miriam Gontijo de – 2018	Moraes, Miriam Gontijo de – 2018			Moraes, Miriam Gontijo de – 2018		
Abordagem etnográfica	Hartel, Jenna - 2010 Quesada, Manuel Alejandro Romero; Quintana, Ania Rosa Hernández – 2015	López, Miriam Tomás – 2004		Morando, María Agustina – 2021			
Abordagem facetada	Ménard, Elaine; Mas, Sabine; Alberts, Inge – 2010				Broughton, Vanda – 2006b		
Análise de conteúdo	Guimarães, José Augusto Chaves; Matos, Daniela Fernanda de Oliveira; Santos, André Ynada dos; Sales, Rodrigo de Sales – 2015 Freitas, Juliana Lazzarotto; Bufrem, Leilah Santiago; Breda, Sônia Maria – 2016; Amorim, Igor Soares; Café, Lígia Maria – 2017; Vital, Luciane Paula; Hernandez, Bianca Ferreira; Santos, Andreia dos – 2019; Café, Luísa Chaves; Cavalcanti, Vanessa Oliveira de Macêdo; Barros, Camila Monteiro de; Vital, Luciane Paula; Alvarez, Edgar Bisset – 2022	Lykke Nielsen, Marianne - 2001 Alencar, Maira Fernandes; Cervantes, Brigida Maria Nogueira – 2019a; Alencar, Maira Fernandes; Cervantes, Brigida Maria Nogueira - 2019b; Fujita, Mariangela Spotti Lopes; Branco, Luciana Beatriz Piovezan Rio; Alves, Roberta Caroline Vesu – 2020; Araújo, Dayanne Albuquerque ; Cavati Sobrinho, Heliomar -2020; Nunes, Flávia Rodrigues Elias; Maculan, Benildes Coura Moreira Dos Santos; Almeida, Mauricio Barcellos – 2021; Nunes, Flávia Rodrigues Elias; Maculan, Benildes Coura Moreira Dos Santos; Almeida, Mauricio Barcellos – 2021; Alencar, Maira Fernandes; Cervantes, Brigida Maria Nogueira – 2021	Colín, Marisela – 2023	Peña, Gilberto; Penã, Catalina – 2015; Lingiardi, Vittorio; Carone, Nicola; Semeraro, Giovanni; Musto, Cataldo; D’Amic, Marilisa; Brena, Silvia – 2020	Guimarães, José Augusto Chaves; Matos, Daniela Fernanda de Oliveira; Santos, André Ynada dos; Sales, Rodrigo de Sales – 2015; Vital, Luciane Paula; Hernandez, Bianca Ferreira; Santos, Andréia dos – 2019; Barite, Mário; Rauch, Mirtha – 2020		

Aspectos metodológicos	Análise de Domínio	Tesauros	Etnia, Raça, Cultura e Léxico	Léxico	Etnia, Raça, Cultura e Análise de domínio	Tesauros e Análise de Domínio	Etnia, Raça, Cultura e Tesauros
Análise do discurso	Barros, Thiago Henrique Bragato – 2023	Marianne Lykke Nielsen - 2001				Garoz, Alexei Zaldúa - 2006	
Análise de citação	Grácio, Maria Cláudia Cabrini; Olveira, Ely Francina Tannuri de – 2013; Guimarães, José Augusto Chaves; Matos, Daniela Fernanda de Oliveira; Santos, André Ynada dos; Sales, Rodrigo de Sales – 2015; Almeida, Josiana Florêncio Vieira Régis de; Dias, Guilherme Ataíde – 2019	Guimarães, José Augusto Chaves; Matos, Daniela Fernanda de Oliveira; Santos, André Ynada dos; Sales, Rodrigo de Sales - 2015 Hjørland, Birger – 2013b					
Análise de cocitação de autores	Moya-Anegón, Félix; Vargas-Quesada, Benjamón; Herrero-Solana, Víctor; Chinchilla-Rodríguez, Zaida; Corera-Elvaz, elena; Munoz-Fernandez, Francisco J. – 2004; Moya-Anegón, Félix; Herrero-Solana, Víctor; Jiménez-Contreras, Evaristo – 2006; Arencibia-Jorge, Ricardo; Vega-Almeida, Rosa Lidia; Martí-Lahera, Yohannis – 2007; Miguel, Sandra; Moya-Anegón, Félix; Herrero-Solana, Víctor – 2008; Grácio, Maria Cláudia Cabrini; Olveira, Ely Francina Tannuri de – 2014; Zhao, Dangzhi; Strotmann, Andreas – 2022; Song, Yanhui; Lei, Lixin; Wu, Lijuan; Chen, Shiji – 2023	Schneider, Jesper W. – 2005; Hjørland, Birger – 2013; Grácio, Maria Cláudia Cabrini - 2016			Wang, Shengang - 2019		

Aspectos metodológicos	Análise de Domínio	Tesauros	Etnia, Raça, Cultura e Léxico	Léxico	Etnia, Raça, Cultura e Análise de domínio	Tesauros e Análise de Domínio	Etnia, Raça, Cultura e Tesauros
Acoplamento bibliográfico	Grácio, Maria Cláudia Cabrini - 2016 Song, Yanhui; Lei, Lixin; Wu, Lijuan; Chen, Shiji - 2023	Hjørland, Birger - 2013					
Análise de co-palavras	Smiraglia, Richard P. - 2015 Hu, Kai; Wu, Huayi; Qi, Kunlun; Yu, Jingmin; Yang, Siluo; Yu, Tianxing; Zheng, Jie; Liu, Bo - 2018	Schneider, Jesper W. - 2005					

Fonte: Elaborado pela autora.

Em “*Managing documents at home for serious leisure: a case study of the hobby of gourmet cooking*” de 2010, Hartel realizou um estudo de caso utilizando a Análise de Domínio para analisar documentos da culinária gourmet, essa atividade se enquadrava como um *hobby* entre profissionais da gastronomia nos seus ambientes domésticos. A pesquisa utilizou a **abordagem etnográfica** combinada à Teoria Fundamentada e com o auxílio do software de análise qualitativa e quantitativa *NVivo*. Percebeu-se que esse artigo envolveu uma abordagem etnográfica viável para a pesquisa em domínios não acadêmicos, ou seja, o foco esteve em um grupo de pessoas e em atividades profissionais e cotidianas, assim seu intuito foi mapear das técnicas culinárias e conhecimento da Gastronomia.

Hartel (2010) explicou que a abordagem etnográfica foi efetivada a partir do contato com 20 cozinheiros gourmet nos Estados Unidos, dos quais foram entrevistados em seus domicílios e, posteriormente, foram realizadas visitas guiadas e inventários nos locais das suas coleções de informação culinária. Conforme notado, essas coleções se caracterizaram enquanto biblioteca pessoal de culinária que variaram em conteúdo (informação culinária), distribuição no espaço e sua função na atividade desses cozinheiros.

As bibliotecas pessoais no *hobby* da culinária gourmet também podem existir em toda a classe mais ampla de hobbies de fazer e consertar (ou artesanato). **Estas atividades (como tricô, fotografia e construção de modelos, entre muitas outras) são apreciadas por dezenas de milhões de pessoas em todo o mundo e exigem habilidade técnica e conhecimento tópico.** O estudo sistemático e a descrição dos fenômenos de informação num contexto de hobby é um primeiro passo para a compreensão de padrões informacionais mais amplos nestes domínios otimistas, prósperos e ricos em informação (Hartel, 2010, p. 869, tradução nossa).

No trecho acima, uma menção ao conhecimento “tópico” e à habilidade técnica nos faz refletir que as pesquisas empíricas realizadas em domínios podem se ater aos interesses e temáticas não científicas, confirmando a noção de domínio entendida por Hjørland como *hobbie* (2002; 2017) e as dimensões e sujeitos do conhecimento Smiraglia (2011, 2015a) que não são exclusivas, em muitos casos, de natureza acadêmica e científica. Nesse caso, as bibliotecas pessoais são espaços possíveis de expressar essa ideia e são locais cuja concentração de documentos, a prática da leitura e da competência informacional se fazem presentes e oportunos para os pesquisadores e profissionais da informação.

O detalhamento da pesquisa etnográfica se caracterizou pelo acumulado de materiais e a gestão de dados dos registros etnográficos, dentre os quais transcrições de entrevistas, notas de campo, descrições, fotografias e memorandos, além de um processo de análise indutiva e a codificação de conceitos e descrições relevantes (Hartel, 2010).

Em outro caso, no artigo da temática de tesouros de 2004, “*Diseño de un tesoro conceptual de folklore de la Región de Murcia*”, López desenvolveu um tesouro de recuperação da informação denominado de TEFOLKMU segundo uma metodologia qualitativa respaldada na abordagem etnográfica (aplicada à área do folclore) e na abordagem experimental típica das pesquisas de natureza linguístico-cognitiva.

López (2004) explicou que o desenvolvimento do tesouro foi obtido com a captação e o tratamento de fontes originais (tradição oral e protocolos notariais) e fontes documentais impressas e inéditas. Assim, subsidiando o esquema conceitual e estrutural dos seguintes macro descritores que não possuíram uma classificação padronizada: trajes; instrumentos musicais; cancionero; intérpretes; comarcas da região de Múrcia; expressão folclórica; brincadeiras infantis e artesanato popular. O dialeto murciano, logo o léxico característico da região, delimitou e apontou os limites tanto geográficos, quanto em termos linguísticos e culturais, uma vez que a abordagem etnográfica os recortes geográficos e históricos são imprescindíveis na execução da pesquisa de campo, além da investigação das fronteiras culturais e linguísticas.

A metodologia de López ocorreu em dois momentos. No primeiro momento o autor empregou técnicas qualitativas aplicadas à observação participante e à captura de fotografias relativas aos objetos e às entidades que dariam nomes as unidades conceituais; prosseguiu com a realização das entrevistas e o estudo da bibliografia e de documentos; e por último a análise de conteúdo e de discurso. O segundo momento é destinado exclusivamente para a elaboração e o *design* do tesouro segundo nas normas UNE 50-106-90 e as recomendações de Aitchison e Gilchrist (1979) e mediante as etapas de apontada por López (2004):

1. Identificação do tema – definição da totalidade e da relação dos temas secundários;
2. Seleção de fontes – extração e seleção de termos da tradição oral, protocolos notariais, manuais de folclore, dicionários, enciclopédias, revistas;
3. Recopilação terminológica – aplicação do método dedutivo e indutivo para identificação dos termos, representação conceitual e estabelecimento das relações lógicas;
4. Normalização dos termos – controle semântico, estrutural e léxico do vocabulário eleito para a pesquisa, atentado para o princípio da univocidade previsto pela TGT;
5. Estrutura conceitual do tesauro;
6. Estabelecimento de relações de sinonímia.
7. Estabelecimento de relações associativas;
8. Implementação tecnológica – aplicação do software *Beat 2.3* no gerenciamento e disponibilização interativa do tesauro;
9. Construção do dicionário terminológico; e
10. Apresentação do tesauro (TEFOLKMUR).

As pesquisas de López em 2004 e Hartel em 2010 apontam para as tipologias documentais e registros como materiais indispensáveis à coleta e à compreensão de aspectos conceituais. Desde documentos textuais, visuais e orais, a abordagem etnográfica pode ser percebida por uma interferência e participação direta do pesquisador, demonstrando a presença da subjetividade e da introdução de questões históricas e inúmeros fatores e fatos de um determinado espaço e tempo. López, 2004, e Hartel, 2010, não se referem à abordagem antropológica e sim à pesquisa etnográfica e ao método etnográfico.

Caracterizado como pesquisa etnográfica provinda dos artigos da temática dos léxicos, Morando em 2021 apresenta: *Los mecanismos de innovación léxica y la reconfiguración de la modernidad indígena en guaraní chaqueño*. A pesquisadora deu ênfase à significação e à construção de processos que englobaram diretamente a identidade oral e a escrita das línguas indígenas.

As dificuldades elencadas e, talvez as barreiras da pesquisa envolvendo léxicos indígenas, perpassam não somente, por empecilhos comuns da ciência e dos mecanismos de objetividade, mas aqui revisitando o texto de Morando, pode-se prospectar outros desafios,

principalmente, ligados às línguas e aos inúmeros aspectos que exigem e atestam os falantes indígenas como os primeiros a serem considerados nessa dinâmica. A pesquisadora argentina busca entender como a inovação e a reconfiguração do léxico indígena da língua chacoana guarani (chiriguano) estão estabelecidos e como se manifestam no cotidiano.

Ainda considerando que essa comunidade indígena enfrenta e sobrevive com suas origens étnicas diante das consequências do cenário e dos elementos da modernidade, dentre os quais aquelas negativas referentes ao desuso de palavras e ao apagamento de suas tradições orais, foram destacadas as seguintes inovações lexicais: “a criação de novas palavras, a incorporação de empréstimos e as operações de extensão semântica pelas quais palavras já existentes na língua adquirem novos significados” (Morando, 2021, p. 173, tradução nossa).

Morando não definiu com clareza como aconteceram os procedimentos etnográficos, contudo declara que documentos, dentre eles dicionários e glossários, foram o foco particular que permitiu, inclusive, um mapeamento dos registros elaborados por missionários durante as missões religiosas e a ocupação dos territórios indígenas. O relato ainda esclarece as influências diferentes etnias participantes da caracterização linguística, além de apresentar a amplitude lexical presente no domínio escolar e acadêmico, jurídico, social e cotidiana através de uma perspectiva antropológica, recuperando os aspectos etno-históricos (Morando, 2021).

Pondera-se nesse sentido, que há uma aproximação com os significados sociais e o olhar sociológico que reimprime nos aspectos teóricos metodologias afins e aptas à caracterização do domínio nos contextos sociais e culturais. Essa justaposição da abordagem com o campo social existe. Contudo, é importante reconhecer que a etnografia se revela e retoma o conceito e as teorias da cultura como parte crucial para as investigações, notando o funcionamento desse método em direção aos temas sociais.

A **abordagem facetada** esteve presente no trabalho de Ménard, Mas e Alberts, em 2010. O artigo foi recuperado na literatura de múltiplas frentes: da Análise de Domínio; da Análise de Domínio e Etnia, Raça e Cultura. As autoras relatam o desenvolvimento de uma classificação facetada, objetivando a elaboração de uma taxonomia para a navegação e recuperação da coleção digital e museológica da Artefacts Canada. A princípio a metodologia buscou aplicar a Análise de Domínio na fase exploratória para compreender o domínio museológico, no tocante das taxonomias museológicas, e para orientar etapas que pudessem arquitetar esse sistema de organização no âmbito das plataformas digitais. Detalhadamente, o objetivo foi delimitar o escopo e a cobertura da proposta taxonômica; explorar e detectar as

potenciais facetas e fornecer recomendações acerca da seleção de facetas (Ménard; Mas; Alberts, 2010).

O tratamento terminológico dos vocabulários dos bancos de dados e interfaces museológicas consistiu em uma etapa de análise essencial que estabeleceu o elo entre a proposta analítica de domínio – vista aqui através das abordagens metodológicas descritas por Hjørland (2002a) enquanto a de Produção de classificações especiais e tesouros e de Estudos terminológicos” e a de LSP (linguagens para fins especiais) e estudos de discursos – com a abordagem e o método analítico sintético proposto por Ranganathan (1937;1967). As autoras não definiram a metodológica enquanto abordagem facetada, porém o material descritivo das seções e o escopo da pesquisa deixam implícito tal denominação e aplicabilidade. Podemos conferir no trecho a seguir:

Três etapas essenciais foram seguidas no desenvolvimento da estrutura de classificação. Primeiro, o domínio da taxonomia foi dividido em conceitos elementares. Esses conceitos foram reunidos em classes fundamentais. Essas classes foram a base a partir da qual as facetas foram selecionadas. Em segundo lugar, as facetas foram organizadas em uma estrutura coerente realizada simultaneamente com a escolha e definição das facetas de nível superior e das suas subfacetas. O número de facetas de nível superior, bem como a profundidade da estrutura taxonômica, foram restritos. O número de facetas de primeiro nível foi limitado a sete e o nível de profundidade reduzido para quatro, desenvolvendo até cinco em alguns casos. Terceiro, o último A etapa do processo de seleção de facetas e subfacetas implicou a normalização das facetas e seus descritores usando fontes confiáveis. As facetas e seus descritores foram desenvolvidos inicialmente em inglês, mantendo o controle dos recursos em francês, quando disponíveis (Ménard; Mas; Alberts, 2010, p. 528-529, tradução nossa).

As etapas descritas demonstram uma adaptação favorável da prática metodológica que utilizou a análise e a categorização em facetas para a projeção estrutural em ambiente de plataformas e sistemas da Web, já que durante décadas o método e a abordagem facetada eram destinados às classificações e aos esquemas manuais e semiautomáticos. Nesse sentido, as autoras se encontram nas observações e propostas de Broughton (2002; 2006a; 2006b) e Vickery (2008) no que concernem a apresentação de unidades terminológicas e vocabulários temáticos em níveis estruturais e semânticos elencados para a recuperação e representação de objetos e entidades.

Há desafios expressos que redefinem a complexidade das unidades utilizadas nesse método. Primeiramente, se levam em consideração a dificuldade de compreensão e a designação de significados, pois os termos e, conseqüentemente, os conceitos utilizados estarão apresentados em níveis cujo controle de dispersões e a univocidade precisam ser aplicadas. As

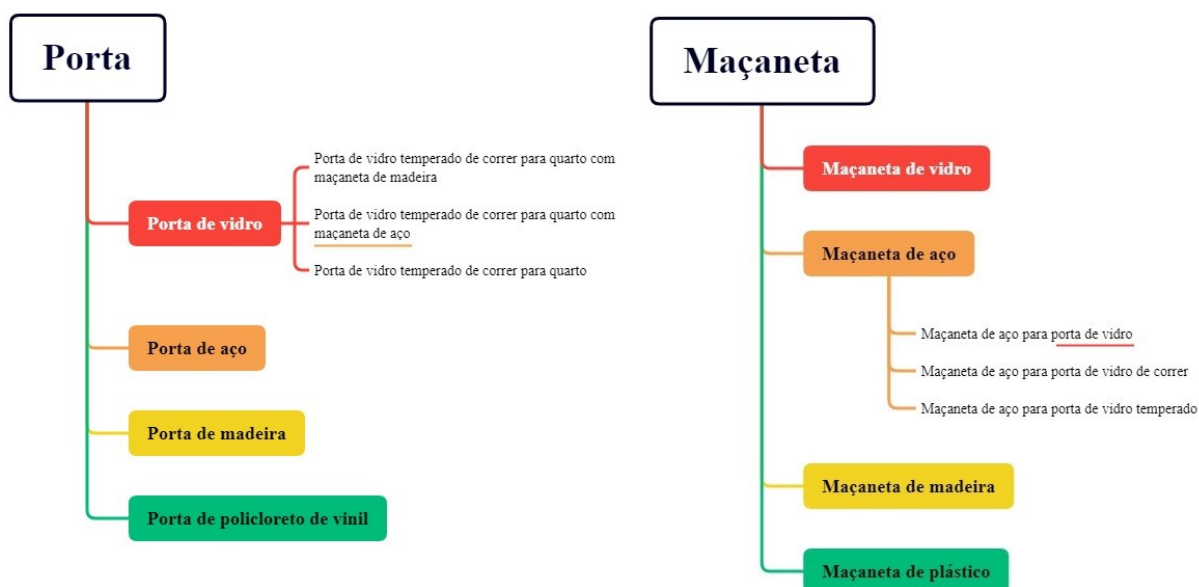
temáticas envolvendo artes e museus, como também, os vocábulos na internet se deparam com uma realidade complexa dos sistemas e páginas da intranet e da internet.

Vickery (2008), por exemplo, afirmou que ao aplicar a proposta facetada, devemos ter em mente a diferenciação entre objetos e entidades, considerando as descrições de teor simbólico presentes na internet e da necessidade de julgar a relevância dos assuntos e sua disposição em facetas organizadas para atender as necessidades dos usuários.

Em “*The need for a faceted methods of information retrieval*” de 2006, Broughton trata da abordagem facetada como meio de estruturar informações em coleções digitais. Dentre os benefícios da aplicação da classificação facetada nos documentos digitais, a autora citou um em especial: a estrutura e metodologia que permitem a conversão desse sistema para as listas de cabeçalhos de assuntos e tesouros (Broughton, 2006b).

Uma característica dessa metodologia descrita no artigo se refere à preferência por uma lógica de indexação de “baixo para cima”. Apesar dessa abordagem implicar em níveis de facetas e suas possíveis decomposições hierárquicas, são nas inúmeras combinações que tornam a experiência dos usuários satisfatória. A partir das explicações da autora, elaborou-se um exemplo dessa abordagem em dois cenários possíveis: o primeiro com termos advindos do léxico do português e o segundo com termos do léxico da língua indígena Apurinã.

Figura 10 – Exemplos facetas em estrutura construída segundo a abordagem facetada



Fonte: Elaborado pela autora.

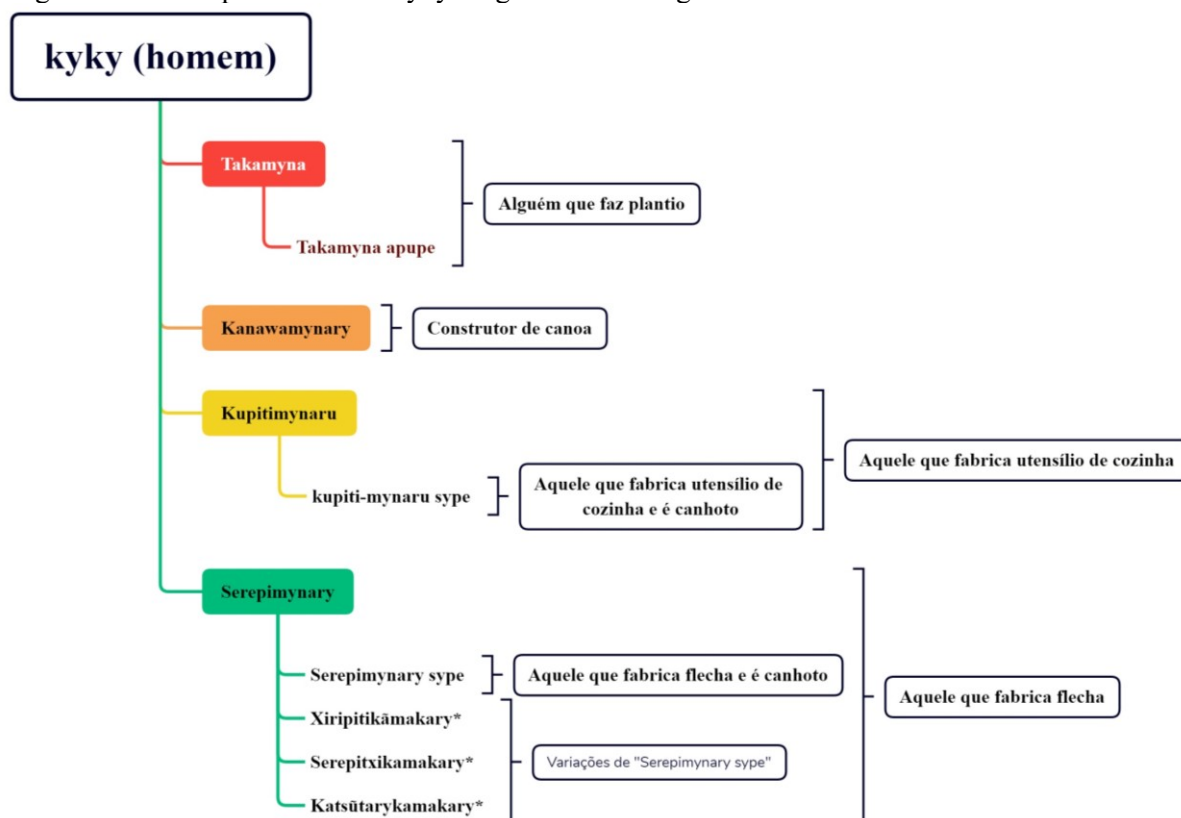
Vejamos na Figura 10 as facetas “Porta” e “Maçaneta”, ambas apresentadas na categoria “matéria” conforme as cinco descritas por Ranganathan (1937; 1967), apresentadas

de acordo com um padrão lógico-semântico cujas possibilidades se estendem para outras combinações possíveis. A ordenação e a seleção das unidades podem ser auxiliadas pelos cânones direcionados à formação dos renques (Exaustividade, Exclusividade e Sequência útil), quando se nota que a unidade conceitual e terminológica “porta” segue outras subfacetas detalhadas em diversas composições e com a proximidade semântica.

Esse exercício da abordagem facetada foi exemplificado por Broughton (2006b) ao citar a faceta “meias cinzas” – meias cinzas > meias de lã cinza > meias de trabalhos de lã cinza; meias de caminhadas de lã cinza; meias de lã cinza para caminhadas; meias de lã cinza até o joelho para caminhadas – em uma estrutura que a autora considerou como taxonômica e não classificatória, já que a classificação em facetas está elencada com a combinação de outras facetas e com a cobertura de um corpus terminológico mais amplo.

Esse exemplo é bem próximo da estrutura de um tesouro como foi o caso do Thesaurofacet. Além disso, levou em consideração às orientações da International Organization for Standardization no padrão 25964-1 e sua ideia de que a análise baseada em facetas “[...]é útil para gerar hierarquias que estejam de acordo com as regras para relacionamentos hierárquicos, conforme estabelecido em 10.2 [relações hierárquicas], porque esses relacionamentos são válidos apenas para conceitos pertencentes à mesma categoria geral” (2011, p.68, tradução nossa). Considerando essas orientações, podemos visualizar na Figura 11 o modelo baseado em unidades da língua indígena Apurinã:

Figura 11 – Exemplo da faceta “kyky” segundo a abordagem facetada



Fonte: Elaborado pela autora a partir da consulta dos verbetes²³.

O conceito “kyky” está representado segundo a categoria fundamental “personalidade”. Esse esquema demonstra alguns aspectos do léxico Apurinã refletido em especificações semânticas e morfológicas diferentes daquelas apresentadas pelo português. A unidade léxica “sype”, apontada em duas subfacetas, corresponde a qualificação da palavra que a antecede, assim sua significação denota “pessoa canhota”. Há ainda outras regras relevantes da formação desse léxico a serem consideradas em paralelo a aplicação da abordagem facetada. São nítidas as possibilidades de classificação e enumeração de termos e conceitos em relação às facetas. Entretanto, é necessário compreender até que ponto os cânones e princípios da facetados podem ser aplicados, principalmente, em léxicos não originários das línguas latinas;

Na leitura em Ménard, Mas e Alberts e Broughton, percebe-se que essa abordagem funciona para tesouros e taxonomias bilíngues e de línguas de origens latinas (românicas), mas há indícios que podem indicar dificuldades em relação a sistemas que utilizam uma língua latina e uma língua indígena da América Latina, pois durante a aplicabilidade é imprescindível fornecer e configurar as facetas conforme os sistemas linguísticos (morfologia, semântica, fonologia) e nem sempre essas línguas possuem características sistemáticas semelhantes.

²³ Verbetes apresentados na macro e microestrutura do Dicionário Apurinã (Padovani, 2020)

Então, é possível que em algum momento não exista equivalente ou correspondência para equiparar todos os conceitos e termos em tesouros e vocábulos bilíngues ou multilíngues. Além disso, o método e a classificação em facetas são exaustivos e específicos, devendo respeitar demandas de um sistema complexo e multifacetado, do qual em alguns casos pode apresentar empecilhos durante a manutenção, as atualizações terminológicas e lexicais durante os ajustes relativos à exaustividade e à especificidade na indexação. Porém, na Figura 11 a abordagem se mostrou aplicável.

Quanto à aplicabilidade, Ménard, Mas e Alberts – 2010- trabalham no primeiro momento com documentos. Contudo, após a aplicação da abordagem de facetas, são realizados testes e a efetivação de um processo de avaliação por parte dos usuários quanto à pertinência da taxonomia, visando atender as necessidades, os critérios de usabilidade e a conformidade desse instrumento para arquitetura da plataforma da Artefacts Canada.

Broughton, em 2006, não apontou detalhadamente as etapas, embora enfatizou um trabalho preliminar na temática, seleção e análise dos termos, para depois dispor e arquitetar as facetas e as subfacetas conforme os procedimentos lógicos e os relacionamentos semânticos. Os dois artigos mencionados acima atentam para o desenvolvimento de taxonomias e tesouros como fortes instrumentos na arquitetura de informação de sites e ambientes da web, como foi apontado por Rosenfeld, Morville e Arango (2015).

Dentre as abordagens diretamente ligadas às teorias, Moraes, em 2018, apresentou a **abordagem epistemográfica** ao colocar em pauta a tensão identitária na Organização do Conhecimento. É importante salientar que esse artigo de Moraes intitulado de “Tensão Identitária e Organização do Conhecimento: olhar epistemográfico” foi recuperado nas seguintes categorias de pesquisa: 1. Análise de Domínio; 2. Etnia, raça, Cultura e Análise de Domínio; 3. Tesouros. A relevância do estudo ainda procede pela exposição dos conflitos epistemológicos, uma vez que a abordagem epistemográfica tem início com o apontamento de conflitos, dentre eles aquelas epistemologias hegemônicas e aquelas epistemologias não evidenciadas ou não incluídas de fato no cenário dos discursos e instituições do conhecimento.

A abordagem epistemográfica surge com Garcia Gutierrez e a partir de seu pensamento de torná-la uma ferramenta na Organização do Conhecimento (Moraes, 2018). As desigualdades epistemológicas existentes em linguagens, classificações e esquemas contribuem com as desigualdades sociais, culturais e as problemas de acesso e o desnivelamento dos conhecimentos (científicos, populares, filosóficos, religiosos). Outro ponto, é reivindicação de que esse pensamento disseminado nos discursos busca dar notoriedade as epistemologias decoloniais e introduz o pluralismo (Garcia Gutierrez, 2006 apud Moraes, 2018). Dessa forma,

A epistemografia tem como objetivo apontar as contradições da epistemologia, identificando os conhecimentos propositalmente excluídos dos processos contemporâneos de inscrição e fluxo, e que são produzidos por comunidades consideradas não epistemológicas. É concebida como uma configuração transdisciplinar que tem como objeto a organização horizontal e interativa dos conhecimentos e, por extensão, da exomemória, entendida como todas as inscrições humanas do conhecimento ou lembranças, em suporte físico ou digital e que também lida com conceitos bem conhecidos e mitificados pela Epistemologia moderna, tais como categoria, classificação, ordenação, método, conhecimento, formalização, metacognição ou representação (Garcia-Gutierrez, 2006 apud Moraes, 2018, p. 5).

Nesse sentido, a epistemografia busca por igualdade de condições entre as tradições das teóricas do conhecimento quando permite que uma perspectiva transdisciplinar e horizontal possa notar e delimitar categorias em um mesmo plano. Essa configuração se mostrou, primeiramente, no entendimento das epistemologias e das teorias que envolvem o objeto de estudo, seja um sistema, como os tesauros, seja um domínio ou assunto científico específico. Posteriormente, analisando Moraes, entendeu-se que, em um segundo momento, essa abordagem desenha e orienta as investigações e a construção de instrumentos do conhecimento.

Moraes (2018) enxergou essa abordagem na compreensão de epistemologias que espelham e promovem a diversidade étnica e racial, dentre as quais através da categoria etnoconhecimento e seu potencial de contribuir com o regaste e as representações dos saberes dos povos originários e suas distintas etnias. Portanto, a epistemografia presente na investigação da autora foi identificada como uma abordagem teórica e que se efetivou nas discussões conceituais e metodológicas acerca das representações (termos, conceitos) utilizadas nas implicações da questão da identidade indígena e das epistemologias feministas. Ou seja, em termos práticos, é uma abordagem que se inicia com reflexões epistemológicas, perpassa por mapeamentos e apontamentos conceituais, para então fomentar uma síntese de tal conhecimento e suas possibilidades.

Em “Abordagem sociocultural da informação” de 2006, Nascimento revisitou e apontou questionamentos acerca do conceito de informação. Esse trabalho, em especial, foi recuperado nas categorias de pesquisa “Tesauros” e “Etnia, raça, cultura e Análise de Domínio”, e buscou através da **abordagem sociocultural**, combinada com os aspectos conceituais da comunidade discursiva e Análise de Domínio, investigar o conceito da informação e as suas implicações para a CI.

Para compreender a abordagem sociocultural recorreremos à exposição da autora em direção às abordagens da CI, comumente, manifestadas através de seus paradigmas. São nos estudos de Hjørland que argumentos são elaborados para justificar que o foco científico incide

na informação e suas dimensões sociais e culturais. O esquema abaixo (Figura 12) representa a distinção entre as abordagens e seus olhares.

Figura 12 – Abordagens da Ciência da Informação em Nascimento (2006)

Abordagens da Ciência da Informação			
Paradigmas	Abordagem	Processos	O olhar
Cognitivo	Indivíduo	Psicológicos	Organização e tratamento da Informação
Físico	Sistema	Tecnológicos	Organização e tratamento da Informação
Social	Domínio	Sociais e culturais	Informação construída

Fonte: Nascimento (2006, p. 31).

Ao revisitar os paradigmas e a Análise de Domínio enquanto um paradigma social, a autora revela que a abordagem direcionada para o domínio envolve processos sociais e culturais. Apesar de não definir de modo objetivo a abordagem sociocultural, alguns pontos destacados na conclusão expuseram em evidência que essa configuração tem:

O resgate das origens ontológicas do conceito informação – dar forma a alguma coisa, permite-nos entender o processo dinâmico de (in-) formação – a essência do objeto. O que foi até aqui exposto orienta-nos para o entendimento da informação pelo olhar sociologizado de Hjørland, onde temos que o estudo das estruturas informacionais pertencentes às comunidades discursivas de um domínio de conhecimento organiza as práticas informacionais construídas historicamente, culturalmente e socialmente. É a possibilidade de provocar a produção de outro conhecimento, não linear, mas circular, que valorize as inter-relações culturais, ambientais, sociais, econômicas e políticas construídas para enfrentar de forma mais coerente e atuante os desafios atuais da sociedade (Nascimento, 2006, p. 33).

A partir das observações da autora, notou-se indícios de que a abordagem sociocultural compartilha das mesmas premissas do paradigma social: o foco em domínios; e a compreensão da informação como um conceito dinâmico pertencente às relações e às esferas sociais e culturais. Essa abordagem concede mudanças e tendências, principalmente, metodológicas, o que podemos considerar que, nesse caso os métodos e os instrumentos, devem ser priorizados os estudos das comunidades e os fatores históricos e culturais como relevantes para o ambiente e os atores da pesquisa. Concluindo desse modo, que estudos étnicos e raciais são excelentes candidatos para tal perspectiva.

Schneider e Borlund (2004) e Schneider (2005) descrevem e introduzem o desenvolvimento dos tesauros utilizando as análises e os fundamentos bibliométricos frente as duas principais **abordagens: as semiautomáticas e as manuais**. Podemos denotar que as abordagens do tipo manual ocorrem exclusivamente por meio trabalho intelectual, sem o processamento por máquina. As abordagens semiautomáticas são aquelas que operam com o auxílio parcial de profissionais durante as etapas e os processos facilitados por metodologias e tecnologias de processamento e tratamento dos dados.

Nesse interim, os conjuntos de métodos e análises bibliométricos se mostraram propostas viáveis para auxiliar nas etapas de elaboração dos tesauros a partir dos modos semiautomático e manual. Embora, a completude e a eficiência não estão ligadas à introdução de mecanismos e ferramentas exclusivamente automáticas para a realização e o cumprimento dos objetivos metodológicos. Há consideráveis obstáculos a serem superados, dentre eles a questão da subjetividade do profissional associada análise da semântica e das propriedades textuais reflexos das línguas naturais.

Durante as leituras, pode-se verificar que as tendências de construção de tesauros em processo e abordagens semiautomáticas tiveram notoriedade, inicialmente, durante meados da década de 1980 e 1990, chegando aos anos 2000 com pesquisas circunscritas em métodos, algoritmos e análises em diferentes combinações e contextos. Duas dessas tendências representam de fato a aplicabilidade de metodologias com pouca interferência humana (semiautomática) ou, ainda, aquelas sem a interferência direta de profissionais (automática) nas principais etapas de elaboração e manutenção de tesauros. Essas foram identificadas em Schneider e Borlund (2004) e Ibekwe-SanJuan (2006), concentrando-se em:

- a) Métodos estatísticos – estabelecidos conforme à associação de co-ocorrência de palavras de primeira ordem;
- b) Métodos que utilizam o contexto linguístico - por meio da associação de co-ocorrência de palavras de segunda ordem.

Esses métodos objetivam, além de extrair e selecionar termos candidatos, associar palavras de acordo com mensuração de suas ocorrências em um dado texto, frase ou unidade textual. Os métodos que utilizam o contexto linguístico para se estabelecerem, associados em métodos e instrumentos estatísticos, embora se diferenciem por enfatizarem o contexto sintático do qual as palavras estão dispostas. Ou seja, a associação de **co-ocorrência de palavras de primeira ordem** (métodos estatísticos) prioriza o agrupamento de palavras próximas e sua

frequência, sem nenhuma ordem específica na colocação sintática do texto, já a associação de **co-ocorrência de palavras de segunda ordem** (contexto linguístico) agrupa palavras em contextos semelhantes ao se valer de critérios linguísticos e da disposição de cada elemento textual (Schneider; Borlund, 2004; Ibekwe-Sanjuan, 2006).

Embora a diferenciação seja expressa, alguns métodos linguísticos fazem uso de medidas estatísticas para possibilitar o tratamento dos termos. Esses se apresentam mais compatíveis a realidade da prática tesimal, uma vez que os critérios linguísticos se aproximam das orientações e diretrizes recentes, principalmente, quanto a facilitar a identificação de sinônimos ou quase sinônimos, algo que na leitura os autores demonstraram mais problemas e dificuldades em relação as associações de primeira ordem.

Nesse sentido, Ibekwe-SanJuan (2006) atestou que a **abordagem linguística** – aquela que ele trata como sinônimo de métodos linguísticos e de ocorrência de segunda ordem – oferece para a construção dos tesouros três abordagens específicas: a abordagem distributiva relativa à similaridade funcional; a evidência sintática interna; e os marcadores de relação semântica procurados em corpus e recursos semânticos internos.

Os métodos de primeira ordem fixados na co-ocorrência de palavras, se comparados com os de segunda ordem, não se submetem a coesão semântica, pois verificam e agrupam palavras próximas umas das outras sem considerar uma harmonia em relação ao eixo e as relações sintáticas. Aqueles de segunda ordem, em oposição, preocupam-se com o contexto, frequentemente, atribuído as palavras e a natureza sintática das frases, sintagmas e orações.

As orientações destinadas às escolhas e à utilização de assuntos temáticos para fins de recuperação da informação, tomando como ponto principal a prática da indexação, despontam como promissoras nas etapas de seleção, tratamento e análise de termos candidatos. Nesse tópico e nos estudos de análise de indexação, Mai, em 2005, se referiu as seguintes abordagens: **abordagem centrada no usuário; abordagem centrada em documento; abordagem orientada a documento; e a abordagem centrada no domínio**. Essa última relativa à Análise de Domínio e aquilo que incumbe no trabalho do indexador a combinações de procedimentos que priorizam o domínio como primeiro passo para desempenhar tal atividade. A partir daí foram identificados outros trabalhos caracterizados por pelo menos uma dessas três abordagens.

A primeira distinção é entender aquela a respeito do documento, ou seja, procedimentos que retomam com exclusividade o documento como fonte de informação para a análise e extração de descritores para indexação ou, no caso dos tesouros, denotam a garantia

literária. Essa garantia considera que o documento é a fonte primeira para executar as etapas iniciais e mantenedoras da estrutura dos termos e de suas implicações. Mai (2005, p. 600, tradução nossa) difere ambas no seguinte trecho:

Na abordagem centrada no documento, o indexador toma o documento como ponto focal para análise, mas ao selecionar os termos do índice, o indexador tem “em mente as questões, tanto quanto podem ser conhecidas, que podem ser colocadas à informação” “sistema” (ISO, 1985, Seção 6.3.3). Essas duas abordagens – orientada a documentos e centrada em documentos – pressupõem que o assunto de um documento pode ser determinado independentemente de qualquer contexto ou uso específico. Eles diferem na seleção dos termos do índice; na abordagem centrada em documentos, os termos do índice são selecionados de acordo com as necessidades e questões dos usuários, enquanto a abordagem orientada a documentos não inclui uma avaliação das necessidades dos usuários.

Como informado essas perspectivas são formalizadas em atividade como a seleção de termos, o que poderia ser estendido para a compilação terminológica baseada nas fontes documentais.

A **abordagem orientada para usuários** retoma o princípio da garantia de uso, ou seja, verificar nas respostas e nas indicações dos sujeitos que utilizam os sistemas para guiar a prática indexical. Verificou essa abordagem em “*Towards a user-oriented thesaurus for non-domain-specific image collections*” de Yoon publicado em 2009. O autor apesar de não definir a abordagem orientada para o usuário deixa claro que a pesquisa esteve fixada na orientação dos usuários de coleções on-line da plataforma Flickr. O autor explorou as tags fornecidas pelos usuários para mapear termos indexadores de imagens, agrupando em categorias (cor, abstrato, específicos, genéricos) e possibilitando a verificação de termos que pudessem viabilizar a apresentação de termos genéricos e específicos em seus relacionamentos.

As marcações sociais entraram em pauta nessa abordagem, do qual a maneira de encaminhar a metodologia recorre a participação dos sujeitos e essa pode ser identificada a partir dos produtores de conhecimento e dos processos constantes de interpretação. Yoon (2009) considerou muito bem a marcação social no desenvolvimento de um sistema orientado ao usuário e verificou como um espaço semântico dado a partir das tags conseguem refletir nas características de documentos imagéticos.

Embora considere essa ideia promissora para a construção dos tesauros, ainda é necessário estudo para corroborar com a eficácia dessa abordagem, visto que há diferenças entre os recursos obtidos a partir das tags e aqueles presentes em um tesouro. Esse estudo em especial, apontou um limiar que poderá atingir métodos e análises aplicados em conjunto de marcadores sociais como ferramentas importantes para incluir termos étnico-raciais em tesauros e vocabulários.

Shiri, Revie e Chowdhury (2002) corroboraram destacando que a perspectiva centrada no usuário foi estabelecida para abordar uma série de questões pouco entendidas e relacionadas com as questões comportamentais e cognitivas, relatando ainda que os sistemas de recuperação da informação suscitavam a preocupação dos especialistas com as maneiras pelas quais os sujeitos pesquisavam seus termos em fórmulas e na expansão de suas consultas. Esses estudos centrados no usuário investigam a seleção e a combinação durante as consultas, sobressaindo fatores implicantes no uso dos termos de um tesouro, o que para título de avaliação poderá implicar em detalhes e indícios de readequação terminológica e de ajustes na política de indexação e na perspectiva de representação dentro dos sistemas de recuperação.

Martín-Moncunill, Sicilia-Urban, García-Barriocanal, Sánchez-Alonso em 2015, motivados pelo grande volume de termos do vocabulário AGROVOC, buscaram avaliar a especificidade dos termos em relação ao domínio. Para alcançar esse objetivo os autores se aproximaram de uma abordagem orientada para usuários ao utilizarem esses sujeitos como especialistas durante a pesquisa, considerando ainda a orientação para os domínios do conhecimento de acordo com Albrechtsen (1995), Hjørland (2002) e Mai (2005). Martín-Moncunill, Sicilia-Urban, García-Barriocanal, Sánchez-Alonso ainda orientaram à noção de abordagem centrada no domínio distribuída nas etapas de indexação.

Hjørland (2013b) apresentou a **análise de citações e acoplamento bibliográfico** enquanto técnicas bibliométricas viáveis para a identificação de termos candidatos dos tesouros, já que ambas notam tipos de parentescos e a distância semântica. Contudo, essas análises não podem ser consideradas isentas de subjetividade, já que os analistas bibliométricos tendem trabalhar em um determinado viés, recorte epistemológico, recorte temporal e com um conjunto de dados de determinada área ou temática.

Verificou-se que parte significativa das publicações se caracterizavam ou buscaram introduzir a **análise de conteúdo** como ferramenta metodológica. Um grupo de autores e publicações utilizaram essa ferramenta para desenvolvimento de seus estudos quanto ao mapeamento de categorias e suas inferências, são eles: Amorim e Café – 2017; Alencar e Cervantes – 2019a, 2019b, 2021; Fujita, Branco e Alves – 2020; Nunes, Maculan e Almeida – 2021; Vital, Hernandez e Santos – 2019; e Rauch e Barité – 2020. A perspectiva mais expressiva nessa série de documentos foi a análise de conteúdo na proposta de Bardin (2011). Essa autora defende que essa análise oferece técnicas sistemáticas para tratamento de indicadores qualitativos, de mensagens e outras comunicações documentadas.

Em “Análise de domínio de políticas institucionais de acesso aberto no Brasil” de 2022, Café, Cavalcanti, Barros, Vital e Alvarez identificaram aspectos fundamentais do acesso

aberto em suas representações (presentes nas políticas institucionais brasileiras) através da Análise de domínio e empregaram uma abordagem qualitativa com o auxílio da análise de conteúdo. Assim, permitiram identificar categorias e termos indicativos dessas políticas institucionais e suas possíveis relações.

Com essa mesma configuração, Guimarães, Matos, Santos e Sales, em 2015, também buscaram reunir referenciais dos congressos ISKO-Brasil, ISKO-Espanha, ISKO-América do Norte e ISKO-França. Os trabalhos recuperados a respeito da Análise de Domínio reforçam essa análise como uma das tendências metodológicas para visualizar domínios, das quais foram apontadas por Hjørland (2002a) e comentadas por Smiraglia (2015a). Esse último teórico evidenciou essa técnica nos estudos de produção de guias, nos estudos epistemológicos e críticos, no desenvolvimento de classificações especiais e tesouros e nos estudos empíricos sobre usuários.

Freitas, Bufrem, Breda, em 2016, corroboraram, significativamente, ao apresentar as opções metodológicas utilizadas na produção científica em Ciência da Informação. Para alcançar os objetivos de identificação e mapeamento dessa produção e de seus autores, as pesquisadoras utilizaram análise de conteúdo para inferir e apontar uma série de fatos e aspectos teóricos-metodológicos ao longo de décadas.

Isso foi imprescindível para verificar que esse modo de enxergar se adapta aos questionamentos da pesquisa como uma visão histórica e possível de observação das variáveis socioculturais, como por exemplo variáveis que indicam a mudança no comportamento político e cultural de determinada comunidade discursiva em detrimento das tradições científicas e das transformações no cenário tecnológico. Apesar de seu foco ainda estar intimamente ligado aos conteúdos de produções científicas, essa técnica, também, pode encontrar entraves ou ser adequada à elaboração e compreensão de categorias bem próxima de discursos não científicos.

Analisar o conteúdo é uma prática bem recorrente na CI, como aponta os resultados das autoras acerca da literatura:

Foram encontradas 367 ocorrências de tipos de análise, com maior incidência para análise documental (168) e análise de conteúdo (136), seguida de análise de tarefas e resolução de problemas (25) e análise de citações (16), porque dispensa a necessidade do entrevistador. Vale ressaltar que a análise do contexto e do discurso são tipos específicos de análise de conteúdo, que aumentaram significativamente após 1982 (Freitas; Bufrem; Breda, 2016, p. 11, tradução nossa).

A presença dessa técnica esteve conferida conjuntamente em estudos bibliométricos, dos quais são caracterizados por normas e meios de produção bem definidos,

tornando mais ágeis as análises se comparadas àquelas pesquisas que utilizam materiais que perpassam da oralidade e dos processos de registro e traduções.

Araújo e Cavati Sobrinho, em 2020, mesclaram uma metodologia composta pelo método integrado para construção de tesouros de Cervantes, análise de conteúdo e os elementos da gamificação e da prototipação, objetivando a finalização da proposta de construção de um aplicativo, o Tabli, destinado à construção de tesouros. Os autores buscaram na análise de conteúdo a possibilidade de categorizar e apontar elementos para o desenvolvimento do aplicativo.

Dois trabalhos acerca do assunto “Léxico” envolvem a análise de conteúdo. No primeiro, Gilberto Peña e Catalina Peña (2015) afirmaram que os fenômenos linguísticos associados à análise do conteúdo dos documentos, dos quais empregados para fins de organização e recuperação, são objetos de estudo bastante revisitados na área de Biblioteconomia e Ciência da Informação. O trecho abaixo se refere aos apontamentos dos autores diante da utilidade das análises de conteúdo e do discurso:

Os diversos tipos de textos e as relações entre eles e a sociedade exibem conexões de diversos tipos de acordo com o campo de onde provêm. A ciência que examina textos está interessada em compreender as propriedades e características comuns no uso da linguagem em todo o espectro de campos que compreendem as ciências sociais e as humanidades. A área de análise e sistematização da informação na Biblioteconomia e Ciência da Informação se manifesta no ato de descrever tipos de textos, dados e conteúdos informacionais por meio dos quais tais textos podem ser localizados nos sistemas. A utilização de processos em comum com outras disciplinas, entretanto, é inegável. Este estudo, aliás, utilizará terminologia e análise lexicográfica (Peña, Gilberto; Peña, Catalina, 2015, p. 26, tradução nossa).

A partir do estudo, percebeu-se que essa técnica contribuiu para reconhecer léxicos e verifica-los de acordo com os discursos dos quais estão situados. Nesse caso específico, a aplicabilidade pode ser acompanhada da extração de termos em um corpus especializados por meio do aporte da Lexicografia. Contudo, a análise lexical e os procedimentos linguísticos e terminológicos estão presentes desde o compilamento e o registro até o agrupamento de unidades.

Em “*Mapping Twitter hate speech towards social and sexual minorities: a lexicon-based approach to semantic content analysis*” de 2020, Lingardi, Carone, Semeraro, Musto, D’Amico e Brena vão ao encontro de Gilberto Peña e Catalina Peña (2015) quando analisaram determinado léxico. Lingardi e colaboradores buscaram o exercício do método da análise de conteúdo voltado para o contexto semântico, do qual os autores chamaram de análise de conteúdo semântica, utilizando um léxico representados por mais de 2 milhões de tuites (frases

e relatos curtos) da plataforma social Twitter para mapear discursos e falas contra grupos minoritários sociais, étnicos, sexuais e de gênero.

Com Colín (2023) analisou-se os nomes pelos quais o racismo é chamado no México, demonstrando uma a relação do seu léxico com fenômenos como o etnicismo, a discriminação e o classismo. A autora agrupou trechos e frases nas seguintes categorias: Existência do racismo; Inexistência do racismo; Classismo; Discriminação; e Terceira Raiz. Apesar do estudo não ter como meta a construção de um vocabulário, podemos entender a possibilidade dessa técnica em conjunto com a perspectiva léxica e o embate discursivo que torna evidentes os discursos e as origens de determinados conceitos complexos e conflituosos para a sociedade.

Em “*Michel Pêcheux's Discourse Analysis: An Approach to Domain Analysis*” de 2023, Barros supôs e buscou provar que a análise do discurso pode colaborar com as análises de domínios. Além de promover a discussão, o autor enfatizou os discursos como elementos da origem dos termos e conceitos, os quais merecem atenção na organização do conhecimento já que cada elemento do corpus textual de cada disciplina ou temática se consolidam a partir do entendimento e da interpretação dos seus discursos.

É notório que a análise do discurso se configura e se revela em algumas vertentes, um exemplo é a frente tradicional dos filósofos e historiadores franceses. Em especial, o autor abordou essa primeira vertente francesa, disseminada por Michel Pêcheux, um dos seus fundadores. A ideia por trás surgiu com a distinção da perspectiva da materialidade histórica e linguística e da materialidade ideológica (Barros, 2023), já que a ideologia se manifesta no discurso e na natureza semântica dele.

Na perspectiva das ciências linguísticas, a análise do discurso permite uma abordagem alternativa para a compreensão dos fenômenos semânticos. A abordagem materialista proposta por Pêcheux na década de 1960 desencadeou uma trajetória frutífera, com contínuas retificações, ajustes, desvios e retomadas. Segundo Pêcheux, o discurso, objeto de estudos em Análise do Discurso, atravessa o caminho do acontecimento, da estrutura e da tensão entre descrição e interpretação da análise do discurso (Barros, 2023, p. 6, tradução nossa).

Nesse sentido, não podemos reduzir a análise do discurso enquanto uma técnica analítica, mas acrescentar a sua noção de campo interdisciplinar e de construção teórica-metodológica típica das ciências sociais. Sua instrumentação perpassa por temas mencionados por Barros, como as lutas de classes, o imaginário e a resistência de determinados grupos

sociais, a partir do momento em que envolve o pesquisador e os materiais linguísticos que disseminam ideologias e fenômenos falados, debatidos e compartilhados.

Desde o seu início, a análise do discurso esteve presente no eixo da análise de domínio e propusemos caminhos baseados na interrelação entre os dois, o que poderia representar um novo método de fazer análises de domínio, especialmente no que chamamos de comunidades epistêmicas ou comunidades discursivas. A terminologia, por exemplo, é fundamentalmente uma teoria da linguística, com aplicabilidade significativa à Ciência da Informação, mas não num campo interdisciplinar como podemos ver na análise do discurso. Assim, sua aplicação deveria ser diferente. Procuramos enfatizar estudos que poderiam ser feitos por meio da análise do discurso: 1) aspectos teórico-metodológicos dos domínios analisados; 2) desenvolvimento de KOS auxiliares para representação de domínios; 3) ética da informação; 4) estudos aplicados a domínios específicos. A análise do discurso tem, fundamentalmente, grande potencial para complementar metodologicamente a área, dada a sua metodologia comprometida ideológica e socialmente e a sua ênfase na linguagem no contexto social (Barros, 2023, p. 6, tradução nossa, grifo nosso).

No trecho, algumas aplicações dentro dos estudos de domínio são citadas como possibilidade de absorver, metodologicamente e teoricamente, a análise do discurso. Para isso, é necessário ter em mente as muitas formas de verificação e interpretação dos discursos. Uma recente, já mencionada aqui, se refere à questão do discurso e suas implicações daquilo que Van Dijk (2010) chama de poder social. Analisando os autores, entendemos que a percepção do discurso se estende aos meios, instrumentos e interesses ideológicos, principalmente no tocante ao controle de representações e quaisquer manifestações semânticas presentes na sociedade. Isso é passível de prática nos discursos ideológicos que imprimem violência e discriminação dos grupos Afrodescendentes e étnicos indígenas da América Latina, como bem mencionados por Munanga (2004) e Quijano (1994).

Com Garoz, em 2006, percebe-se o acompanhar os antecedentes dos estudos da análise do discurso e, também, compreender a definição apontada por Van Dijk:

O próprio Van Dijk afirma que AD [análise do discurso] é: “o estudo do uso real da linguagem por falantes reais em situações reais”, que ocorre em dois níveis: no nível sintático – nos próprios enunciados – e no nível semântico – que envolve o sentido –, e isso permite “reduzir a multidimensionalidade do fenômeno discursivo”; bem como reduzir qualitativamente os aspectos mais significativos. Esses dois níveis de análise são interdependentes, assim como a sintaxe e a semântica. A AD permite conhecer conceitualmente o emissor e, como método, permite isolar e classificar as noções pelas quais determinado conhecimento é exposto, e isso pode gerar uma rede semântica que permite a identificação dos interesses e intenções dos produtores (Garoz, 2006, p. 3, tradução nossa).

Esse trecho explica, em linhas gerais, como ocorre a análise diante dos níveis semânticos e sintáticos. Desse modo, a ordem metodológica confere ao plano linguístico e

lexical níveis que denotam os preceitos técnicos para influir a identificação da forma e do conteúdo. Já a perspectiva salientada por Barros (2023) revelou a origem e a manifestação simbólica e social, uma vez que a distinção entre a materialidade linguística, materialidade histórica e a materialidade ideológica transforma os estudos do discurso em uma maneira de perceber as nuances em torno do que é de fato um discurso, o modo como é comunicado e os seus objetivos diante de tantos assuntos científicos e midiáticos (televisão, rede social).

Em “*A framework for work task based thesaurus design*” de 2001, Nielsen examinou um método misto, fruto da combinação entre a análise do discurso e a análise de conteúdo destinadas à elaboração de tesouros orientados para domínios. O trabalho foi realizado por meio da análise de entrevistas aplicadas em grupos, do qual a autora verificou o conteúdo expresso em frases e conceitos, explanando as opiniões e demandas dos usuários diante do domínio observado. A autora ainda citou as unidades terminológicas e os termos técnicos.

Notou-se características que diferenciam a análise de discurso da análise de conteúdo, a primeira prioriza o contexto de produção e disseminação das mensagens mais do que a segunda. Assim no discurso, o conteúdo se mostra como parte integrante do ato do pensamento, além de ser crucial na articulação e na aquisição de sentidos envolvem preceitos e condições para além daquelas linguísticas, ou seja, convenções e representações políticas, midiáticas, socioculturais etc.

As análises e estudos bibliométricos apontados no referencial teórico da Análise de Domínio, a partir de Hjørland e Albrechtsen (1995), Hjørland (2002) e Smiraglia (2015a; 2015b), estiveram presentes nos documentos recuperados e selecionados. Foram destacadas as análises relacionais dos tipos de citação (de autores, de referências) e cocitação (de autores, de referências), de co-palavras (Co-word) e acoplamento bibliográfico, além das suas representações em redes e matrizes.

Os dados fornecidos da produção disseminada nos meios formais de comunicação científica transitam entre os elementos analíticos. As citações consistem em um desses elementos do tipo relacionais para mensurar e avaliar diferentes fatores e variáveis de desenvolvimento, estagnação ou evolução de uma disciplina ou campo científico. Essa unidade tem uma representatividade significativa por indicar discursos científicos no tocante dos seus consensos, dissensos e conformidades. Grácio e Oliveira entende que:

Uma citação é tomada como um indicador objetivo e claro da comunicação científica, que permite a identificação de grupos de cientistas e suas publicações, com a finalidade de evidenciar os pesquisadores de maior impacto de uma área, apontando seus paradigmas, procedimentos metodológicos pertinentes, bem como os

pesquisadores de “vanguarda” que constroem o novo conhecimento na área (2013, p. 199).

Essa unidade é objeto de outras análises bibliométricas e tidas como influentes, também, nos estudos cientométricos e infométricos. Sua preponderância está ligada ao fato de as organizações e instituições científicas nutrirem a necessidade de monitorar a produção e a disseminação científica. Por outro lado, apesar da hegemonia de frentes epistemológicas e grupos científicos, são por meio das citações que os movimentos de contraposta e surgimento de novas epistemologias podem subsidiar o fortalecimento dos grupos científicos não hegemônicos e representantes de frentes decoloniais (epistemologias latino-americanas e africanas). Mesmo ciente do “efeito Mateus na Ciência” e das instâncias que ainda prevalecem com sua configuração elitista, imperialista, colonial e machista, é possível que as citações sejam também utilizadas como representações das novas e esquecidas epistemologias.

Hjørland (2013b) pontou a análise de citações como uma das análises da Bibliometria que seguem uma abordagem social e dinâmica do conjunto metodológico da Organização do Conhecimento. Na leitura, inferiu-se que o autor considerou a análise de citação como forma de reconhecimento das estruturas conceituais e de visibilidade que, sobretudo, pode ser canalizada em outros procedimentos de mapeamento de sintagmas e discursos. A partir desses procedimentos, entendeu-se a sociabilização do conhecimento em observações essenciais para transformar os produtos da prática da Organização do Conhecimento (gestão, sistemas, instrumentos e tecnologias) adequadas aos usuários e os processos de comunicação científica.

Existem maneiras diferentes de explorar as relações de citação e as relações semânticas hoje. Avram, Caragea e Dumitrache (2012) sugeriram uma melhoria na bibliometria através da introdução da ponderação do valor da citação (usando um grau de similaridade semântica). Tal abordagem pressupõe que a similaridade estatística entre documentos pode ser tomada como uma medida de relacionamento entre assuntos (sem uma discussão desta suposição). Que dois documentos podem estar conceitualmente relacionados, embora não sejam estatisticamente semelhantes, é fácil de demonstrar considerando dois documentos sobre o mesmo assunto escritos em línguas diferentes.¹³ É importante enfatizar que do ponto de vista epistemológico as coisas não são apenas semelhantes: Documentos (e qualquer outra coisa) são semelhantes em certos aspectos e diferentes em outros (Hjørland 2013b, p. 1316, tradução nossa).

Nesse trecho, o autor coloca as relações de citações e relações semânticas no mesmo plano exploratório, tendo em vista que as relações estabelecidas por essas mensagens podem indicar possíveis relações semântica entre os documentos. É através de uma série de métodos, leis e axiomas espelhados em modelos e teorias matemáticas e estatísticas que análises

bibliométricas, principalmente, as de citações, atingem aos objetivos de mensuração, mapeamento e avaliação de variáveis e achados empíricos. Em vista disso, Smiraglia (2015a) esclareceu que a análise de citação está na categoria de metodologias empíricas da Análise de Domínio, acrescentando que as próprias citações fornecem evidências residuais da semântica social de um domínio acadêmico e podem ser classificadas e quantificadas para diferentes utilidades.

Por conseguinte, Guimarães, Matos, Santos e Sales, em 2015, combinaram a análise de conteúdo com a análise de citações para examinar a questão conceitual do domínio da Organização do Conhecimento a partir dos referenciais apresentados em reuniões dos eventos brasileiro, espanhol, norte-americano e francês da ISKO. Os autores verificaram através das citações quais pesquisadores foram amplamente mencionados e referenciados, dentre eles: Hjørland; Dahlberg; Gardin; Mai; Olson; Barité e Thelleffssen (Guimarães et al., 2015). Esse mapeamento e avaliação dos autores expressivos na Organização do Conhecimento e, consequente, sua comunidade discursiva, confirmaram uma forte tendência das análises de citações e as demais análises bibliométricas como aportes metodológicos mencionados por Almeida e Dias em 2019, no artigo “Estado da arte sobre Análise de Domínio no campo da Ciência da Informação brasileira”.

Grácio e Oliveira (2013) explicaram que as análises de citação tiveram um início promissor através dos índices de citação criados por Eugene Garfield e nos trabalhos de Solla Price. O mecanismo de mensuração acontece com a visualização e categorização de indicadores que representam variáveis e dados quanti e qualitativos fornecidos por índices de bases de dados e plataformas de dados científicos. O acompanhamento das citações contribui com a identificação das frentes de pesquisa, grupos acadêmicos científicos, tendências de pesquisas; no monitoramento de índices de impacto e produção entre periódicos e instituições; e permite a elaboração de rankings de desempenho de instituições científicas e o acompanhamento de pesquisas.

Na leitura, verificou-se que a análise de cocitação de autores (ACA) é uma tipologia do aporte bibliométrico utilizado para averiguar as regularidades e repetições das citações estabelecidas entre as relações sociocognitivas dos pesquisadores. Essas são derivadas da análise de citações e abordam as frequências com que dois autores ou documentos são citados de modo conjunto entre os documentos de uma área. Em termos metodológicos, essa citação abaixo sintetiza como acontece essa análise na prática:

Quatro passos principais em uma análise de cocitação de autores são apresentados por McCain (1990). Os dois primeiros são: construção de uma matriz com as frequências de cocitação entre os autores e a conversão desta matriz em uma matriz de índice de cocitação relativa (ou normalizada), entre eles o índice de Jaccard e o cosseno de Salton. Desse modo, os estudos de cocitação de autores baseiam-se em matrizes de valores (absolutos ou relativos) que mensuram a coocorrência entre os autores citados. São matrizes de proximidade, isto é, matrizes que indicam quão próximos ou quão distantes estão dois autores, sob o olhar dos pesquisadores citantes (Grácio; Oliveira, 2013, p. 200).

A cocitação nessa aplicação contribui para reconhecimento das frentes epistemológicas e grupos de autores que costumam escrever sobre o mesmo assunto. Ela é aplicada diretamente em referências bibliográficas e observa a relação entre dois artigos de acordo com o número de publicações que aparecem citadas concomitantemente. Entretanto, a ACA visa a relação entre dois autores de acordo com a frequência com que aparecem citados por um terceiro autor. Ainda aquele tipo de cocitação aplicada exclusivamente nas referências de uma determinada disciplina, periódico ou conjunto pré-estabelecido.

Zhao e Strotmann, em 2022, prosseguiram seus estudos históricos e bibliométricos ao desvendar a estrutura intelectual da Ciência da Informação durante os anos de 2006 até 2020, utilizando a análise de cocitação de autores. Essa forma de verificar concentrou-se na visibilidade dos autores e suas relações cognitivas e sociais, o que foram admitidas pela identificação da movimentação dos grupos de pesquisa, da colaboração entre instituições, da genealogia acadêmica e da verificação do predomínio dos discursos e frentes teóricas. A imagem e a identidade dos autores e das citações são resultantes analíticas que inferem na representação ou no espelho de um domínio ou grupo de cientistas.

Contudo, é em Grácio e Oliveira (2014) que se conclui a importância dos estudos de análise de cocitação de autores em associar os dois índices, absolutos e relativos, para visualização e compreensão das estruturas subjacentes de um domínio científico. Essas associações são estruturadas em matrizes e, geralmente, possuem valores absolutos (independente da ordem em que aparecem). Os indicadores, durante a avaliação dos índices, podem ser normalizados em casos comparativos, pois entre áreas e disciplinas são comuns as diferenças entre os dados e as particularidades de cada indicador (Grácio, 2020; Gläzel, 2003). Para normalização dos indicadores alguns modelos e equações são aplicados, como por exemplo o Cosseno de Salton. O cosseno de Salton calcula o índice normalizado com base nos números da matriz e representa com efetividade a intensidade do relacionamento entre os autores (Grácio; Oliveira, 2013).

Em resumo, Miguel, Moya-Anegón, Herrero-Solana em 2008; Grácio, Oliveira em 2013 e 2014; Grácio em 2016 e Song, Lei, Wu, Chen em 2023 utilizaram **a análise de cocitação**

de autores como análise fundamental para reconhecer a estrutura intelectual e a comunidade que integra um dado domínio. Com a leitura, foi possível notar uma consolidação gradativa dessa análise nos estudos bibliométricos e infométricos, a partir dos estudos de Howard White e Katherine McCain; Howard White e Belver Griffith; e Henry Small. Ambas as pesquisas introduzem as técnicas métricas para viabilizam o descobrimento de comunidades e seus percursos na história de disciplinas e ciências, daquelas questões salientadas na Análise de Domínio e nas suas origens sociológicas do conhecimento, principalmente, em decorrência de acontecimentos e observações entendidas a luz da Sociologia do Conhecimento e da Sociologia da Ciência.

Grácio e Oliveira apontaram como a técnica aplicada na cocitação utiliza métodos e ferramentas para a visualização das estruturas:

A importância dos dados oriundos das matrizes de cocitação e de índice de cocitação relativa, para a visualização e compreensão das estruturas subjacentes de uma comunidade científica, tem dado origem a uma significativa literatura na área da Ciência da Informação. Alguns autores, como Ahlgren, Harneving e Rousseau (2003) e Eck e Waltman (2008), posicionam-se favoravelmente ao uso do cosseno de Salton; outros, como Leydesdorff (2008), posicionam-se pelo índice de Jaccard como medida mais representativa de similaridade para normalização nos estudos de análise de cocitação de autores (Grácio; Oliveira, 2013, p. 202).

Nesse trecho, as autoras optam pela aplicação do cosseno de Salton e do índice de Jaccard para mensurar a proximidade entre os autores cocitados, assim permitiu a leitura dos indicativos de força e ligação existentes nas relações. Outros métodos e algoritmos objetivam a visualização desses relacionamentos, gerando redes de cocitação e redes de acoplamento bibliográfico, de modo a produzir imagens bidimensional e, em alguns aplicativos e softwares, tridimensional.

Moya-Anegón, Herrero-Solana e Jiménez-Contreras (2006) reforçaram o uso da ACA e a análise de cocitação de periódicos (JCA) na identificação de domínios temáticos ao propor uma abordagem conexionista e multivariada associada ao algoritmo neural do mapa auto-organizado (SOM), ao método de agrupamento de Ward e ao escalonamento multidimensional (MDS). Os autores compararam esse caso com as pesquisas desenvolvidas por Howard White e Katherine McCain. Essa metodologia empregada potencializou a descoberta tanto de uma espécie de eixo temático do domínio em questão (no caso a Biblioteconomia e sua relação com a CI), quanto as relações advindas da institucionalização das escolas e agremiações científicas.

Em “*The Intellectual Landscape of the Domain of Culture and Ethics in Knowledge Organization: An Analysis of Influential Authors and Works*” de 2019, Wang utilizou a análise

de cocitação de autores como técnica para conduzir um estudo de Análise de Domínio da cultura e da ética na Organização do Conhecimento. Esse foi o único exemplo que incluiu uma aplicabilidade mais próxima a atenção dada para a temática “cultura”, contudo ainda consiste em uma ausência de estudos dedicados aos temas étnico-raciais e culturais.

Moya-Anegón, Vargas-Quesada, Herrero-Solana, Chinchilla-Rodriguez, Corera-Elvazer, Munoz-Fernandez, em 2004, propuseram uma técnica para classificação de classes e categorias utilizando a análise de cocitação de autores. Assim, reconheceu-se a aplicabilidade dessa técnica através da ocorrência dos assuntos mapeados entre os autores, isso auxilia a verificação da semântica social e o consenso das comunidades discursivas (Smiraglia, 2011, 2012, 2015a), além de material para verificar elementos que descrevem as dimensões social e epistemológica (Hjørland; Hartel, 2003). Já em 2007, Arencibia-Jorge, Vega-Almeida, Martí-Lahera utilizaram a ACA combinada à análise semântica para obter o mapeamento de palavras mais frequentes do domínio da Hiperplasia Prostática Benigna (HPB) com o intuito de serem articuladas na elaboração de uma linguagem estruturada.

Song, Lei, Wu e Chen (2023) compararam a análise de cocitação de autores com o acoplamento bibliográfico de autores, do qual acoplamento se mostrou uma análise mais sensível a observação de autores com alta colaboração institucional, apresentando uma imagem mais clara das estruturas e tendências intelectuais se comparado com a ACA. Assim, o acoplamento foca no elo entre dois artigos a partir da ocorrência de outras publicações citadas em comum entre esses dois artigos.

Contudo, notou-se que Song, Lei, Wu e Chen, em 2023, concordam com Grácio, em 2016, no tocante do uso combinado da análise de cocitação de autores e do acoplamento bibliográfico para gerir uma investigação mais abrangente acerca da estrutura intelectual dos campos de pesquisa. Nessa perspectiva, pode-se considerar que as análises citadas demonstram um mapeamento da literatura de um assunto e seus principais autores, o que adquire o papel de um instrumento para indicar fontes que subsidiam a garantia literária e o corpus terminológico utilizados no desenvolvimento e manutenção dos tesouros.

Ainda na comparação das duas análises, Hjørland (2013b) sugeriu que ambos são bons indicadores de relacionamento entre assuntos, contudo o autor menciona Small (1973) ao referir que as duas análises forneciam padrões significativamente diferentes, sendo o acoplamento bibliográfico menos confiável para medir a similaridade de assuntos do que a cocitação.

Na leitura de Song, Lei, Wu e Chen, notou-se que acoplamento pode acontecer na relação combinada de autores, assim como, entre artigos, medindo a força de atração entre eles

com base nas citações em comum. Grácio (2016) colaborou mencionando que o acoplamento identifica a sobreposição da identidade teórico e/ou metodológica dos pesquisadores, assim como concebe as estruturas (social, cognitiva) na perspectiva deles próprios.

A escolha de qual análise bibliométrica será aplicada no estudo de um domínio depende do objetivo da pesquisa e da configuração do objeto. Se o objeto é um grupo epistêmico e a meta é verificar a trajetória de determinados autores e sua relação com seus filiados, a análise de cocitação de autores é uma excelente opção metodológica.

Contudo, essas análises bibliométricas mencionadas acontecem a partir do fornecimento de dados bibliométricos e indicativos controlados e geridos por bases de dados e diretórios científicos, impossibilitando aplicação em dados científicos não inseridos nessas plataformas ou dados de outros setores do conhecimento não institucionalizados. Limitando assim, a aplicação para alguns grupos e temáticas, ou seja, no caso de domínio do tipo *hobbys* e interesses não acadêmicos, mencionados por Hjørland (2017). Assim, não são viáveis em todos os domínios e algumas circunstâncias requerem gestão e aplicação de outros métodos e ferramentas para tratamento dos dados e documentos a serem analisados.

Hu e colaboradores em 2018 aplicaram um modelo de distribuição de palavras usando aprendizagem no mapeamento semântico. Nesse estudo da análise de palavras na perspectiva bibliométrica, os autores propõem a aplicação do método Term Frequency-Keywords Active Index (TF-KAI), do qual se mostrou uma opção para mapeamento de temáticas em domínios, recorrendo aos instrumentos estatísticos e automatizados e frisando na necessidade de solucionar imprecisões e problemas desencadeados durante a identificação de sinônimos específicos. Essas questões também são ressaltadas a partir dos esforços do Processamento de Linguagem Natural (PLN). Assim como, a preocupação com a metodologia de identificação e encadeamento de sinônimos são argumentados:

Os processos manuais que mesclam sinônimos geralmente fornecem resultados de alta qualidade, mas o processamento manual, entretanto, requer certo conhecimento especializado no domínio. Além disso, o processamento manual é, em muitos casos, complicado quando a quantidade de literatura é elevada. Os métodos automáticos existentes para mesclar sinônimos na análise de palavras-chave usam bancos de dados de conhecimento estáticos e não levam em conta contextos dinâmicos de palavras-chave que podem afetar o significado semântico das palavras-chave. Portanto, é necessário um método para lidar e representar o significado dinâmico das palavras em contextos específicos de domínio (Hu et al., 2018, p. 1033, tradução nossa).

A combinação de métodos automáticos e não automáticos parece no centro da exposição dessa análise, sabendo que é algo recorrente na literatura de tesouros, principalmente, em eleger termos e palavras expressivas do assunto dos documentos. Hu e colaboradores, apesar

de tratar o tema mencionando em linhas gerais da análise de palavras, denotaram a análise de co-palavras como uma versão desta última, mas associada às redes de conexões entre palavras, de modo que elas próprias são “nos” para ilustrar o elo e os níveis de ligação entre documentos e autores.

Em relação a automatização, Schneider, em 2005, investigou a aplicabilidade da análise de co-palavras para construção semiautomática de tesouros, associada às análises de cocitação e de citação. Esse autor pontuou as características dessa análise se referindo em como o foco principal dessa está na proximidade semântica e o quanto ela é tão necessária para realocação de termos diante das dificuldades lançadas frente aos quase-sinônimos e a preservação da significação de sintagmas.

A representatividade das palavras-chave está concentrada no uso de regularidades estatísticas, sintaxe, gramática ou características baseadas em rede (Hu et al., 2018). Assim como nas análises bibliométricas, nesse caso, a **análise de co-palavras** pode ser retratada no estabelecimento de redes ao agregar palavras próximas para a formação de *clusters*. Smiraglia, em “*Ethics in knowledge organization: two conferences point to a new core in the domain*” de 2015, aponta o trabalho sobre a ética na Organização do Conhecimento a partir de termos mapeados na cocitação de autores e na ocorrência de palavras-chave.

Diante das análises bibliométricas, verificou-se alguns aspectos práticos mencionados acima a partir do corpus bibliográfico recuperados na base Web of Science no índice SciELO Citation Index acerca do termo “Tupinambá” ao que se refere a uma das etnias dos povos tradicionais brasileiros de mesmo nome e ou língua. A pesquisa simples foi realizada somente com o termo “tupinambá” como assunto tópico, sendo recuperados 43 artigos, dos quais foram selecionados somente 42 publicações se referiam à temática da língua e do povo Tupinambá, excluindo artigos que tratavam do sobrenome, local e outros sentidos atribuídos a esse substantivo. Para a execução das análises, utilizou-se o software de licença livre *Vosviewer*²⁴, cujo foco está em construir e visualizar redes bibliométricas e fornecer aplicações para o tratamento dos dados.

Primeiramente, a análise de cocitação de autores surgiu segundo os dados tabulados compostos por todos os metadados dos artigos e suas referências completas contidas na extensão de arquivo de texto txt. A matriz de relacionamento foi estabelecida no próprio software. Contudo, essa matriz pode ser elaborada em arquivo de tabela conforme a identificação dos dados de ocorrência como destacou Grácio e Oliveira (2013). A Figura 13 é

²⁴O software criado e atualizado pelo Centre for Science and Technology Studies da Leiden University da Holanda. Disponível em: <https://www.vosviewer.com/>.

um recorte da tabela da matriz dos trabalhos citantes e autores citados elaborada no software Excel.

Figura 13 – Recorte do exemplo da matriz quadrada do total de trabalhos citantes e autores citados da língua e do povo Tupinambá

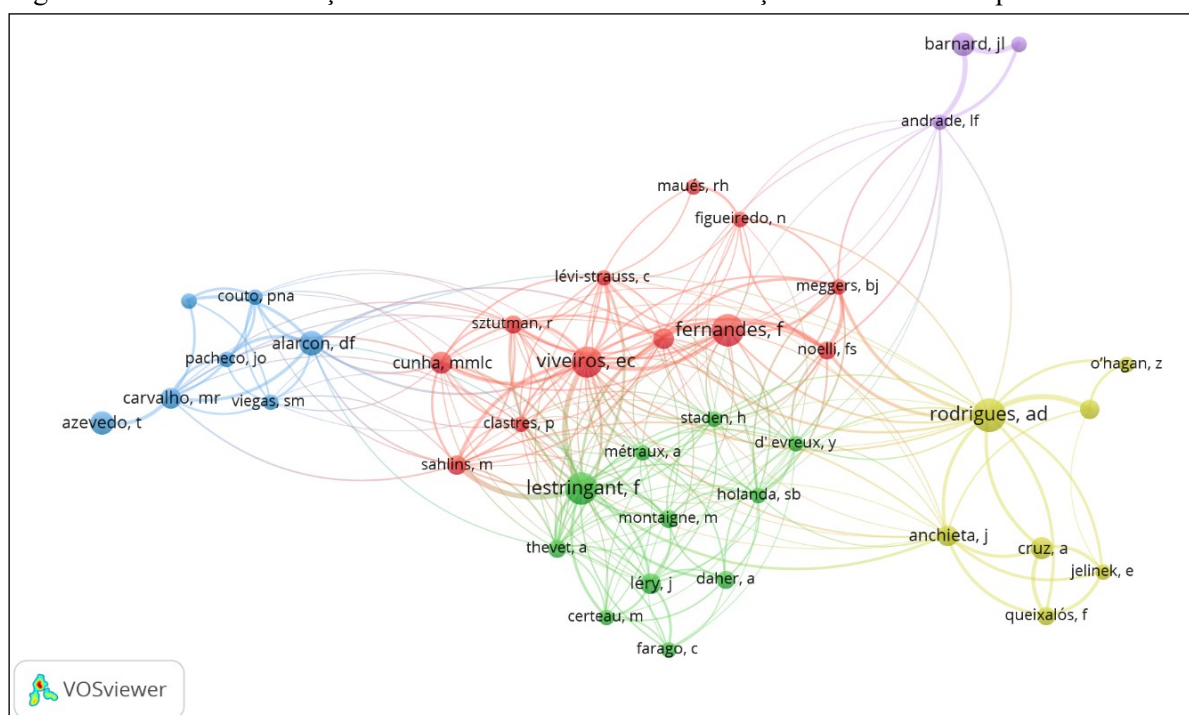
	A	B	C	D	E	F	G	H	I
1		Abbeville Claude	[Agnolin, .	[Alarcon, I	[Albuquerque	[Almeida, [	[Andrade, [	[Bonillo, F	[Chicanga
2	Abbeville Claude, Histoire de l	0	0	0	0	0	0	0	0
3	[Agnolin, Adone] USP	0	0	1	1	1	1	1	1
4	[Alarcon, Daniela Fernandes] U	0	1	0	1	1	1	1	1
5	[Albuquerque, Maria Betânia E	0	1	1	0	1	1	1	1
6	[Almeida, Miguel Vale de] ISC	0	1	1	1	0	1	1	1
7	[Andrade, Luiz F.] Universidad	0	1	1	1	1	0	1	1
8	[Bonillo, Pablo Ibáñez] Univers	0	1	1	1	1	1	0	1
9	[Chicangana-Bayona, Yobenj A	0	1	1	1	1	1	1	0
10	[Chicangana-Bayona, Yobenj A	0	1	1	1	1	1	1	1

Fonte: Dados da pesquisa.

A tabela pontua a ocorrência para cada citante e citados, de modo que os valores inteiros são atribuídos para a quantidade de ocorrência da citação. Os dados da tabela não implicaram valores absolutos e frequência após a normalização dos dados. Seja por Cosseno de Salton, Índice de Jaccard ou Coeficiente de Pearson, ambos para padronizar os dados de modo alinhados ao conjunto dos índices de similaridade e intensidade com que os autores e documentos estão relacionados. Posteriormente, com a visualização dos dados em tabelas, verificou-se a necessidade de padronização dos 803 nomes de autores, assim foi utilizada novamente a tabela do Excel para elaborar o tesauro de autores. Em outros termos, foram realizadas manualmente a adequação dos nomes para sobrenome e iniciais, assim como a eliminação de duplicados ou correção da grafia.

Após o tratamento dos dados dos autores e da importação de todos os arquivos (arquivo de tabulação e arquivo de autores normalizados) para o *Vosviewer*, foram geradas as redes resultantes das análises. A rede de cocitação de autores (Figura 14) se deu mediante aos seguintes parâmetros metodológicos combinados no software: o mínimo de 5 citações por cada autor, sendo que 40 autores foram relacionados por atender os critérios dessas 5 citações; o método full counting (contagem completa dos dados); a normalização do tipo *Association strength* (baseado na força de atração de acordo com o número e a densidade de citações); e os valores de repulsão (-1) e atração (2).

Figura 14 – Rede de cocitação de autores com no mínimo 5 citações do descritor Tupinambá



Fonte: Elaborado pela autora.

A rede gerada segundo a análise de cocitação de autores apresentou a formação de 5 *clusters*, agrupando 40 pesquisadores que estiveram ligados na rede e dispostos conforme cada citação referenciada por um terceiro autor. Os círculos referentes a cada autor apontam uma variação de tamanho, segundo a intensidade com que cada um foi citado. Cada *cluster* representa autores que foram citados em conjuntos, ou seja, o trabalho que cita “Fernandes, F” (em vermelho) também cita “Viveiro, EC” (em vermelho). Os relacionamentos recebem atenção na Tabela 2:

Tabela 2 – autores mais citados conforme a análise de cocitação de autores do termo descritor Tupinambá

Autor	Quantidade de citações	Força entre o total de links
Rodrigues, AD	21	172
Fernandes, F	20	106
Lestringant, F	20	224
Viveiros, Ec	18	219
Alarcon, Df	11	101
Azevedo, T	10	20
Barnard, JL	10	73
Cruz, A	9	69
Cunha, MMLC	9	81
Anchieta, J	8	91
Carvalho, MR	8	86

Fausto, C	8	111
Léry, J	8	58
Cabral, ASAC	7	53
Sahlins, M	7	118
Wosiacki, WB	7	0
Daher, A	6	36
Montaigne, M	6	48
Noelli, FS	6	98
Queixalós, F	6	64

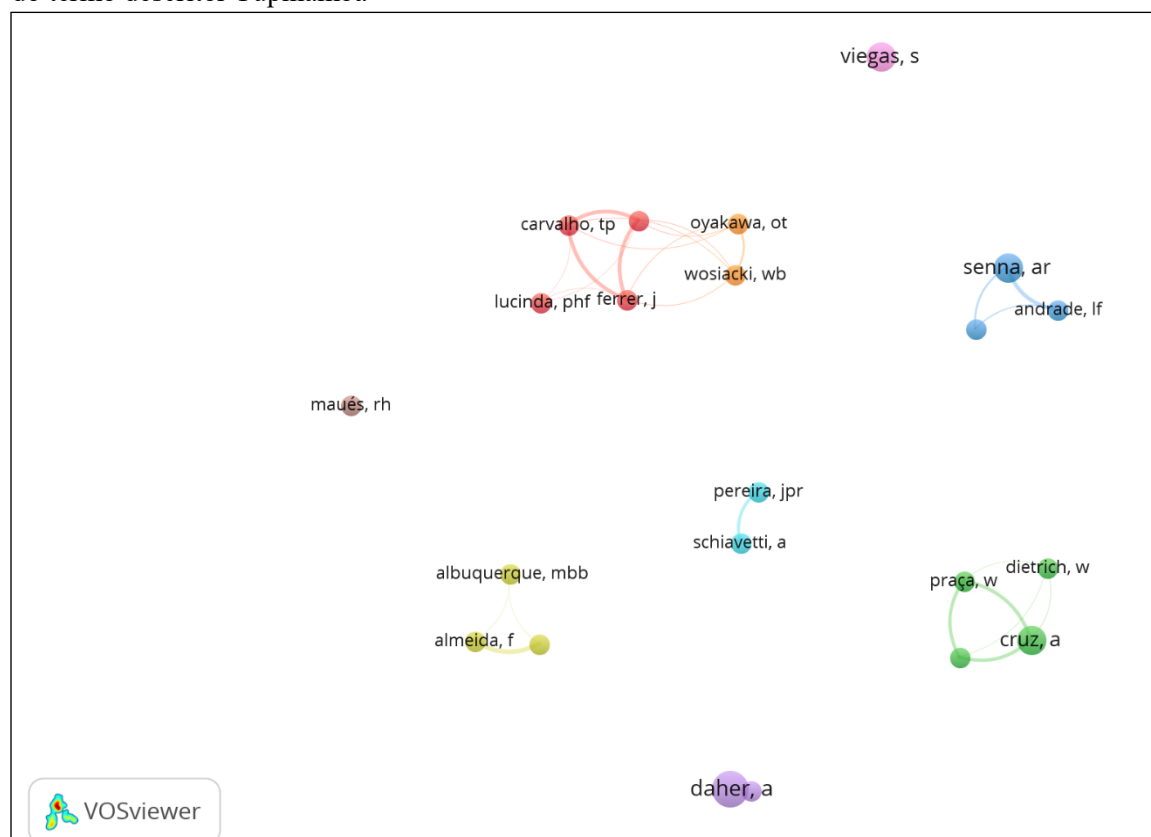
Fonte: Elaborado pela autora.

A Tabela 2 apresentou destaque para o linguista “Rodrigues, Aryon Dall'Igna” ao recebe o maior número de citações, totalizando 21 citações e 172 links estabelecidos, seguido do sociólogo “Fernandes, Florestan” com 20 citações e força de 106 links de relacionamento. O linguista “Queixalós, Francisco” aparece na vigésima posição com 6 citações e 64 links. De acordo com a verificação dos títulos e assuntos das referências, esses pesquisadores mais citados abordam temáticas envolvendo os estudos das línguas étnicas brasileiras, além das culturas, das trajetórias históricas e sociológicas dessas comunidades, cujas particularidades das temáticas dos povos e da língua “Tupinambá” estão inseridas.

Os autores agrupados no *clusters* representam os grupos epistêmicos, é nesse ponto que a análise de cocitação de autores fornece a capacidade de identificação desses grupos tão influentes para a delimitação dos representantes e fornecedores dos discursos e corpus terminológicos, podendo assim, serem relevantes para atesta a garantia literária, filosófica e estrutural. Outro ponto está em reportar quais pesquisadores predominam, isso serviu para nomear quais epistemologias estão exercendo credibilidade nos discursos de um determinado domínio, apontando para o Glänzel (2003); Zhou, Chen e Wang (2009); Oliveira e Grácio, (2013); e Grácio (2016, 2020) sobre as análises bibliométricas consistirem como facilitadoras na compreensão e na visualização de domínios.

Contudo, é por meio do acoplamento bibliográfico que podemos reiterar o que em comum esses teóricos compartilham. A Figura 15 é a imagem da rede de relacionamento dada a partir da análise de acoplamento, da qual considerou no mínimo de 1 citação por autor e 1 documento por autor.

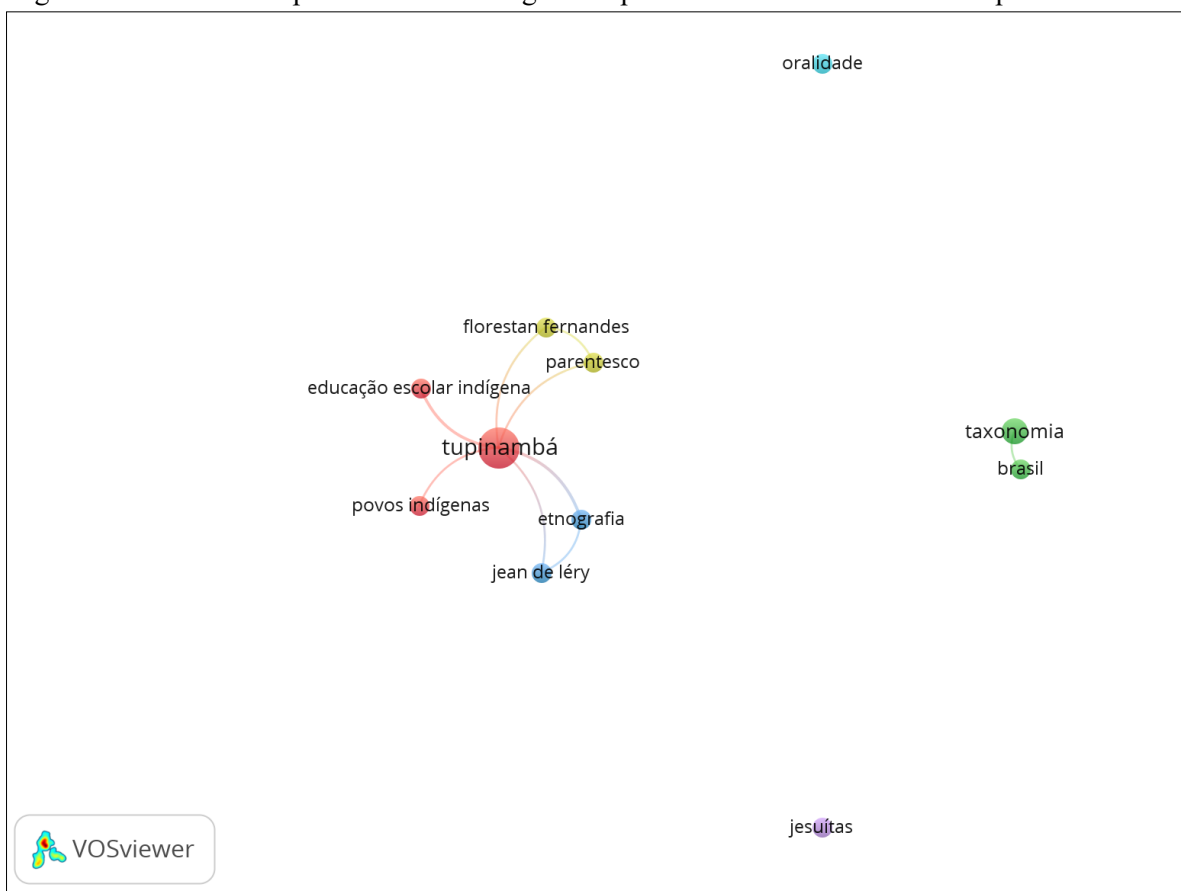
Figura 15 – Rede de acoplamento bibliográfico de autores com no mínimo 1 citação e 1 documento do termo descritor Tupinambá



Fonte: Elaborado pela autora.

Dos 51 autores possíveis de serem acoplados, somente 22 atenderam aos limites de 1 citação e 1 documento. Na rede (Figura 15), podemos conferir que dos 9 *clusters*, 2 deles estão conectados e os demais dispostos sem nenhum relacionamento. Os *clusters* em vermelho (Carvalho, TP; Donin, LM; Ferrer, J; e Lucinda, PHF) e laranja (Cruz, A; Dietrich, W; Magalhães, M; e Praça, W) foram aqueles que compartilharam similaridades, esses ainda podem consistir em vertentes teóricas ou colégios invisíveis. Nesse sentido de investigar esses grupos científicos, a rede de co-palavras (Figura 16) mapeou os 10 termos mais significativos e suas ligações de acordo com os 42 artigos analisados.

Figura 16 – Rede de co-palavras dos 42 artigos recuperados com termo descritor Tupinambá



Fonte: Elaborado pela autora.

A rede de co-palavras considerou as palavras-chave dos autores que ocorreram pelo menos em 2 artigos. Assim, foram realocados 11 termos em 6 *clusters*. O *cluster* destacado em vermelho corresponde aqueles mais influentes, apontando que os termos Tupinambá, povos indígenas e educação escolar indígena mantêm uma vinculação semântica com essa literatura e com a possível ou as possíveis comunidades discursivas. Esses termos mais frequentes são ligados aos *clusters* em amarelo (parentesco, Florestan Fernandes) e azul (Jean de Lery, etnografia).

Os demais clusters não apresentaram relacionamentos, pois algumas palavras-chave da rede foram atribuídas a único artigo ou não ocorreram em outra publicação. Em resumo, o mapeamento das palavras-chave pontua uma parcela expressiva de termos, explicitando a observação de Smiraglia (2015a) de que a análise de co-palavras elucida aspectos qualitativos e envolve uma compreensão muito detalhada do próprio domínio; e a metodologia utilizada por Schneider (2005) para a seleção de termos na construção semiautomática de tesouros.

A seguir métodos e propostas metodológicas foram mapeados no Quadro 11:

Quadro 11 – Métodos e propostas metodológicas

Métodos e propostas metodológicas	Análise de Domínio	Tesauros	Etnia, Raça, Cultura e Léxico	Léxico	Etnia, Raça, Cultura e Análise de domínio	Tesauros e Análise de Domínio	Etnia, Raça, Cultura e Tesauros
Método etnográfico	Quesada, Manuel Alejandro Romero; Quintana, Ania Rosa Hernández - 2015				Romero Quesada, Manuel Alejandro; Quintana, Ania Rosa Hernández - 2015		
Método analítico-sintético							
Métodos de desambiguação semântica, sintática e desambiguação de termos filosóficos		Tiberi, Melissa; Mazzocchi, Fulvio - 2007					
Metodologia de mineração de texto	Lee, Jae Yun; Kim, Heejung; Kim, Pan Jun - 2010						
Metodologia para a elucidação de KOS	Graf, Ann M - 2020						
Metodologia para construção de vocabulários controlados – CoVoMe		Tomaszok, Dominik - 2022					
Metodologias <i>top-down</i> ou <i>bottom-up</i>	Barité, Mario; Fernandez-Molina, Juan Carlos; Guimarães, Jose Augusto Chaves; Ernesto de Moraes, Joao Batista - 2010						
Metodologia híbrida utilizando a Teoria da Classificação Facetada e a Teoria dos Níveis Integrativos	Rodrigues, Anderson Luiz Cardoso - 2015				Rodrigues, Anderson Luiz Cardoso - 2015		
Metodologia de análise fatorial	Pinto, Maria - 2008						
Metodologia combinando PNL superficial com algoritmo de agrupamento e interface de visualização de informações		Ibekwe-SanJuan, Fidelia - 2006					

Métodos e propostas metodológicas	Análise de Domínio	Tesauros	Etnia, Raça, Cultura e Léxico	Léxico	Etnia, Raça, Cultura e Análise de domínio	Tesauros e Análise de Domínio	Etnia, Raça, Cultura e Tesauros
Metodologia holística de construção de tesauros da arte no contexto ocidental		Chen, Sophy Shu-Jiun - 2021					Chen, Sophy Shu-Jiun - 2021
Metodologias no estabelecimento da interoperabilidade em KOS		Zeng, Marcia Lei; Chan, Lois Mai - 2004					
Metodologia de base bibliométrica		Schneider, Jesper W. - 2005					
Metodologias de transformação de vocabulário em SKOS para a Web,		Bezares, Gonzalo Mochon; Rodriguez, Eva Maria Mendez; La Fuente, Gema Bueno de - 2020					
Metodologia orientada para tarefas de trabalho utilizada no desenvolvimento do design de tesouro		Sanatjoo, Azam - 2010					
Metodologia para a identificação de sentimento em léxico				Rigo, Sandro José; Alves, Isa Mara da Rosa; Barbosa, Jorge Luis Victória - 2015			
Metodologia de construção de micro tesouro		Reis, Mateus da Silva; Barros, Thiago Henrique Bragato; Santos Junior, Roberto Lopes dos - 2021					
Metodologia de elaboração de tesouro conceitual		Campos, Maria Luiza Almeida; Gomes, Hagar Espanha - 2006					
Metodologia híbrida (Teoria da Classificação Facetada, de Ranganathan e a Teoria dos Níveis Integrativos) na elaboração de uma estrutura de classificação	Rodrigues, Anderson Luiz Cardoso - 2015				Rodrigues, Anderson Luiz Cardoso - 2015		
Metodologia para desenvolvimento de léxico terminológico				Pajić, Vesna et al. - 2018			

Métodos e propostas metodológicas	Análise de Domínio	Tesauros	Etnia, Raça, Cultura e Léxico	Léxico	Etnia, Raça, Cultura e Análise de domínio	Tesauros e Análise de Domínio	Etnia, Raça, Cultura e Tesauros
Metodologia de observação detalhada de grupos sociais utilizando léxico				Bonaccorsi, Andrea; Chiarello, Filippo; Fantoni, Gualtiero - 2021			
Metodologia estatística para extrair subjetividade				Sarvabhotla, Kiran; Pingali, Prasad; Varma, Vasudeva - 2011			
Metodologia de disponibilidade lexical combinada com a teoria dos grafos				Blanco Correa, Óscar Elías; Salcedo Lagos, Pedro; Kotz Grabole, Gabriela Emilce - 2020			
Metodologia combinando os estudos de tradução baseados em corpus; o da linguística de corpus; e a terminologia				Paiva, Paula Tavares Pinto; Camargo, Diva Cardoso de; Xatara, Cláudia Maria - 2008			

Fonte: Elaborado pela autora.

A priori a análise dos métodos utilizados e fornecidos como propostas deparou-se com divergências terminológicas, principalmente, quanto aos métodos e suas diferenças metodológicas se comparados com as abordagens. Schneider e Borlund (2004; 2006), Schneider (2005) mencionaram os métodos bibliométricos enquanto abordagens bibliométricas, contudo levando em consideração o contexto de construção e manutenção semiautomática de tesauros. É notório na leitura desses teóricos que a equivalência atribuída para métodos e abordagens está direcionada às opções analíticas, aos procedimentos e às ferramentas estatísticas e linguísticas utilizadas de forma combinada ou não para obter variáveis, indicadores e unidades terminológicas a serem comparadas e mensuradas quantitativa e qualitativamente.

Os métodos bibliométricos identificados foram: a análise de citação, análise de contexto de citação e análise de co-palavras (Schneider; Borlund, 2002, 2004; Schneider, 2005, 2006), ambos discutidos anteriormente. Wang (2019) tratou da análise de citação de autores (ACA) como método, entretanto o autor aplica a mesma no contexto de mapeamento dos autores e identificação da representação intelectual de um domínio. Apesar de Wang não utilizar

a ACA para o desenvolvimento de tesouros, o autor esteve próximo das colocações de Schneider; Borlund (2002, 2004) e de Schneider (2005, 2006) quando reconhece o potencial métrico destinado aos estudos empíricos em disciplinas, áreas e temáticas de pesquisa. Concordou-se com Schneider (2006) quanto ao salientar que os métodos bibliométricos contribuem para a abordagem semiautomática de elaboração e continuidade dos tesouros. Contudo, esses métodos não são suficientes e requer outras opções metodológicas e perspectivas, incluindo aquelas manuais e de caráter intelectual.

O **método etnográfico** é outro método tratado, anteriormente, enquanto abordagem. Quesada e Quintana (2015) apontam pontos convergentes desse método e a Análise de domínio, sendo considerados pressupostos socioculturais adaptados as bases interdisciplinares e transdisciplinares na CI. O autor apontou de modo objetivo a questão epistemológica, a ambientação social da pesquisa e os procedimentos formais que caracterizam essa prática.

Discutir o método etnográfico e sua relação teórico-conceitual com a Análise de Domínio requer eleger os aspectos sócio-culturais em comunidades discursivas como algo que a priori interfere diretamente em quais caminhos e limites metodológicos são fornecidos em tese. Os autores destacam que o ponto de convergência se dá por compreender que esse método se aproxima da descoberta dos domínios de acordo com suas perspectivas sociopsicológicas, sociolinguísticas e sociocognitivas conferidas na abordagem social (Quesada; Quintana, 2015). Isso permite analisar as comunidades e o desvelamento de suas bases socioculturais, como anteriormente, foram destacadas por Hjørland e Albrechtsen (1995).

O Quadro 12 é uma síntese da apresentação do método por Quesada e Quintana.

Quadro 12 – Etapas e elementos do Método Etnográfico em Quesada e Quintana (2015)

Processos	Descrição
Primeira fase	Delimitação e seleção do objeto de estudo; delineamento dos objetivos e das questões de pesquisa.
Segunda fase	Processo empírico-analítico caracterizado pela pesquisa de campo, coleta e análise dos dados.
Técnicas	Descrição
Observação participante	Técnica primária que consiste no exercício da reflexividade do pesquisador diante do objeto estudado. O objetivo está nas ações de registro e descrição do que foi observado, esclarecendo que o pesquisador está mediante aos seus sentidos participando e buscando perceber acontecimento e captar todas as informações. Apresenta níveis de participação: passiva (com a presença no cenário observado, com participação limitada); moderada (participação com pouca interferência); ativa (busca interagir e realizar aprendizagem no cenário); completa (participação completa do pesquisador em todas as atividades e pessoas que compõem o grupo).

Entrevistas qualitativas e aprofundadas	Técnica de captação de discursos, falas e saberes orais dos participantes.
Discussão em grupo	Contato com sujeitos e atores que compõe o cenário, sendo essencial para compreender as representações sociais das comunidades e buscar compreender consensos. Promove a participação e a interação dos pesquisadores e dos sujeitos da pesquisa.
Ferramentas	Descrição
Diário e caderno de campo	Instrumento para descrição e registro de observações frutos das reflexões e sensações captadas pelo pesquisador.
Ficha de inscrição	Registro dos participantes, permitindo coletar dados relevantes e possíveis variáveis da pesquisa, com idade, região, escolaridade etc.
Análise e escrita etnográfica	Descrição
Reflexão analítica acerca dos dados	Tarefa de reflexão do pesquisador, constituindo de uma ação fundamental para a organização e seleção dos dados.
Seleção e redução de dados	Permite a síntese para melhor entendimento dados e posterior organização.
Organização e categorização dos dados	Mapeamento de categorias e verificação de elementos relevantes que possam agrupar os dados.

Fonte: Adaptado de Quesada e Quintana (2015)

Na leitura dos elementos, é possível averiguar que Quesada e Quintana apontaram características desse método semelhantes àquelas exploradas por Hartel (2010) e López (2004), pois esses autores estiveram diante de um campo empírico cuja dinâmica dos atores sociais precisaram ser captadas, já que há complexidade relativa ao limite estabelecido entre pesquisador e os participantes. Além disso, ambos têm os documentos e as tipologias documentais como fontes primeiras de suas pesquisas, e em segundo lugar a tradição oral expressa nas entrevistas.

As metodologias relatadas se referem à sistematização de uma série de elementos prescritivos e abordagens. As principais destacadas são aquelas que integram contribuições da Ciência da Informação e aquelas fornecidas pelos estudos de análise e processamento de linguagem natural e léxicos. Nesse interim, Lee, Heenjung Kim e Jun Kim (2010) revelaram que a **mineração de texto**, também conhecida na infometria como tomografia de banco e dados, propõe a extração automática de sintagmas nominais em conjuntos textuais. Os autores tomaram como basilar três momentos metodológico: o primeiro para evidenciar os perfis dos descritores e dos conjuntos de documentos (periódico), gerando uma matriz de similaridade conforme os vetores e os pesos atribuído às unidades (descritores individuais e conjunto

documental) e seus relacionamentos; no segundo momento, os agrupamentos são estabelecidos a partir das unidades próximas; e por último, a disposição das redes em *clusters*.

Lee, Heenjung Kim e Jun Kim (2010) ainda utilizaram o algoritmo *Pathfinder Network* (PFNet), o método paralelo de agrupamento de vizinhos mais próximos²⁵, a geração da matriz de similaridade de perfil e a análise e mapeamento multidimensional. Nesse sentido, inferiu-se que essa proposta está estabelecida em processos automatizados e que confere pouca interferência humana. Contudo, a atividade intelectual é percebida nos procedimentos conclusivos de avaliação do domínio em questão. Aqui os autores obtiveram resultados significativos no domínio das bibliotecas digitais, produzindo um plano comparativo das unidades analíticas e das tendências reveladas. Novamente nesse estudo, percebe-se que a metodologia consegue ser efetivada em conjunto com avaliação intelectual dos profissionais e os métodos que organizam os documentos e fornecem metadados estruturados.

Em casos dos quais a literatura e o fornecimento de conteúdo não são geridos e organizados em bancos e diretórios de dados científicos e/ ou bem estruturados, as metodologias utilizadas em redes sociais em ambientes virtuais ou métodos tradicionais, como os da abordagem etnográfica, podem ser fortes candidatos em pesquisas cujo dados ou materiais não são estruturados.

Tomamos como exemplo, a **metodologia desenvolvida para a identificação dos sentimentos expressos nas mensagens textuais em linguagem natural** aplicada por Rigo, Alves e Barbosa (2015) que utilizou a abordagem composta pelas teorias da Semântica Lexical Computacional com base nos estudos de análise de corpus, na descrição de unidades lexicais e na implementação de sistema computacional para identificar a representação do léxico da emoção. Diante da proposta desses autores, entendeu-se que o léxico dos usuários – o foco de Rigo, Alves e Barbosa foi estudantes do ensino à distância – podem conter, entre outras representações relevantes, a representação dos sentimentos.

A partir daí o léxico investigado pela Semântica Lexical Computacional e pelo processo analítico de sentimentos possibilita a identificação, além de unidades terminológicas, das outras nuances da linguagem natural. Considerando os tesouros, esses aspectos elucidados denunciam pontos convergentes, principalmente aqueles conceitos e terminologias confrontadas pelas representações e léxicos característicos da diversidade de gênero, como os estudos.

²⁵ O método paralelo de agrupamento de vizinhos mais próximos citado pelos autores foi seguido da sigla PNNC.

Graf (2021) apresentou uma **metodologia de elucidação de sistemas de organização do conhecimento** caracterizada pela Análise de Domínio aplicada aos sites e portais de galerias artísticas do Grafite – ou grafitte, compreendido enquanto expressão artística ou arte de rua. O conjunto metodológico utilizado pela autora foi adequado às condições e à disposição das imagens fotografadas por diferentes grupos e artistas do Grafite, das quais estavam disseminados na web. Isso leva em consideração como o pesquisador nota a organização nesses espaços, principalmente, entendendo que as informações, vocabulários ou termos associados a esses documentos visuais, ou ao próprio discurso e às falas de seus criadores. Assim, apresentando perspectivas e formas diferentes daquelas tradicionalmente tratadas sob as lentes das teorias da classificação facetada e bibliográficas e a Teoria do Conceito de Dahlberg.

Os passos fornecidos por Graf (2021) possibilitaram não apenas uma taxonomia dos termos da Arte do Grafite e do vocabulário tesauroal, mas também permitiram promover a conscientização de aspectos para descrição e a disponibilização de um esquema de metadados e perfil de aplicação dados. Assim, aproximou a metodologia da autora com estruturação de vocabulários e tesouros no contexto da Arquitetura da Informação da web apresentada por Rosenfeld, Morville e Arango (2015) e, apesar desses autores reafirmarem os procedimentos de rotulagem de ambientes diferente daquele apresentado por Graf, ambos se encontram no tocante das etapas iniciais de entendimento da temática ou do domínio que dita e correspondem as demandas de representação em ambientes virtuais. Isso está amplamente demarcado por terminologias de diferentes níveis de especialidades, como também, vocabulários fixados na língua de seus usuários.

Analizando esse trabalho das imagens do Grafite, notou-se que o primeiro passo empregado foi concebido no mapeamento das informações dos sites e na identificação dos rótulos. Logo, percebeu-se uma estrutura organizacional do vocabulário utilizado para dividir uma série de características e classes conceituais, desde estilos acerca dos movimentos artísticos até mesmo os nomes das técnicas de criação e os locais onde os grafites foram elaborados. Em segundo, foi possível articular categorias que realocaram diferentes tipos de características e formas de organização dessas imagens. Assim, observou-se que a autora aplicou métodos de categorização com o auxílio de software qualitativo e de organização de dados, além de elucidar as seções que agruparam diferentes temáticas.

O padrão de conteúdo de dados de Catalogação de Objetos Culturais da *Visual Resources Association*, bem como rótulos associados aos textos dos documentos visuais e às imagens foram levados em consideração por Graf (2021). Diante desses achados, a categorização

favoreceu o estudo, o que indicou ser uma análise de conteúdo ou análise de categorias. Resumindo, a autora registrou manualmente informações dos sites e vocabulários, posteriormente categorizou e rotulou os vocábulos e conjuntos textuais com o intuito de analisar quanti e qualitativamente. Essa experiência associada à análise crítica do domínio se configura como uma opção favorável para diferentes tipos de estudos, dentre os quais aqueles que não são de temáticas e assuntos acadêmicos, como exemplos as fontes de informação sobre as etnias e os marcadores sociais.

Em um contexto mais recente, Tomaszuk, em 2022, apresentou o CoVoMe como uma **metodologia para construção de vocabulários controlados projetada para diferentes idiomas**. Essa metodologia vai ao encontro dos requisitos e das orientações destacadas nos padrões ISO 25964-1 e 25964-2, pois esclarece de modo objetivo etapas fundamentais e imprescindíveis para a disposição e a composição das relações semânticas, além da estruturação das etapas aproveitadas em sistemas que utilizam, fornecem ou criam SKOS.

Nesse sentido, as etapas metodológicas apontadas no CoVoMe por Tomaszuk (2022) são: Atividade de elucidação de domínio e escopo; Atividade de escolha do tipo de currículo ou documentação – listas de termos, classificações, categorias, listas de relacionamentos; Atividades destinadas a estratégia de tratamento dos conceitos e esquemas conceituais – com base na denominação e na atribuição de definições de acordo interpretação dos significados das unidades; Atividades de tratamento lexical das entidades – com base em rótulos, termos e notações; Atividades de coleções – identificação das coleções ordenadas e não ordenadas – conforme os agrupamentos dos conceitos e suas características ordenadas ou não; Atividades de integração de propriedades de similaridades, hierarquias e associações – relacionamentos semânticos; e por último as atividades de documentação – focadas na apresentação dos termos e das unidades em sua estrutura – e avaliação – processo que engloba a verificação e análise da utilização do vocabulário.

Um dos pontos de destaque dessa proposta de Tomaszuk está na integração das abordagens *bottom-up* (dos conceitos mais específicos para os gerais), *top-down* (dos conceitos mais genéricos para os específicos) e *middle-out* (a partir de conceitos centrais) presentes enquanto princípios que guiam essa metodologia. Nessa integração, foram notados os termos das buscas realizadas pelos usuários. Nesse sentido, verificou-se um avanço em relação aos esquemas construídos e fornecidos pela abordagem *top-down*.

Barité, Molina, Guimarães e Moraes, em 2010, não relataram uma metodologia única a ser seguida, mas relacionaram às **metodologias *top-down* ou *bottom-up*** com os princípios da garantia literária, o que foi possível verificar a importância das decisões em

relação às práticas e às fontes literárias e terminológicas como o conjunto de documentos de especialidades para reconhecimento de conceitos genéricos e específicos.

Chen (2021) sugeriu uma **metodologia holística de construção de tesauro** para permitir que a arte ocidental seja acessível aos usuários chineses e a arte chinesa seja acessível aos usuários ocidentais através das seguintes tarefas: 1) coleta de termos, (2) análise semântica, (3) mapeamento de equivalência, (4) pesquisa sobre construção de sistemas de organização do conhecimento, (5) consultoria desempenhada por grupos de especialistas, (6) construção do conceito, (7) tradução, (8) revisão, (9) desambiguação e (10) revisão e aprovação do grupo de especialistas. Notou-se que a metodologia foi capaz de integrar métodos que vão ao encontro da interoperabilidade em sistemas de organização, incluindo procedimentos de tradução e a tentativa manter a consistência conceitual em ambos os idiomas.

Além disso, há um esforço da metodologia em preservar aspectos da arte em termos culturais e próprios de sua origem. Para os procedimentos de tradução o projeto pode ser complexo, mas com a atenção que a autora dá para a configuração linguística podem suprir as demandas mais gerais, como aquelas de identificação de conceitos equivalentes ou quase sinônimos para tratar de termos sem traduções.

Duas propostas são exclusivas e pensadas para tesouros: **“Proposta metodológica para a elaboração do micro tesauro do acervo de medalhística do Museu do Instituto Histórico e Geográfico do Pará”**; e **“Metodologia de Elaboração de Tesauro Conceitual: a categorização como princípio norteador”**. A primeira publicação, apesar de indicar uma proposta metodológica, apresentou de modo geral três etapas: seleção de documentos; delimitação de um vocabulário especializado, baseado nesses materiais bibliográficos; e a construção de micro tesauro conforme a estruturação de termos e conceitos selecionados (Reis; Barros; Santos Júnior, 2021). Essa articulação não foi detalhada, embora adaptou-se bem à realidade do museu, pois o vocabulário controlado representou termos referentes às medalhas (réalias) com base no mapeamento de palavras selecionadas em documentos sobre esses objetos históricos.

A segunda já buscou reintegrar uma fundamentação pautada na categorização de termos de acordo a Teoria do conceito e método facetado. Nesse sentido, esta última permitiu perceber a realidade empírica diante do objeto que está representado conceitualmente, recordando que o método de faceta é relevante para tal prática, além do método dedutivo fornecer as bases para arquitetar grandes classes estruturadas em cadeias e renques (Campos; Gomes, 2006).

Quanto à aplicação exclusiva em domínios, Rodrigues (2015) promoveu uma metodologia híbrida, utilizando a Teoria da Classificação Facetada, de Ranganathan, e a Teoria dos Níveis Integrativo para descobrir, conceitualmente, o domínio da cultura. O entendimento dessa harmonia de teorias começa quando o autor utiliza teoria facetada para identificar classes gerais (categorias) da cultura. Esse passo demonstrou uma análise da literatura e a seleção de conceitos que se sobressaem nas discussões. Em um segundo momento, notou-se que a Teoria dos Níveis Integrativos introduziu aspectos filosóficos e um modo sistemático de visualizar a construção e a história de tal domínio, apostando na observação do universo que está inserido tal conhecimento em todas as entidades e propriedades ligadas a ele. A proposta realça um estudo teórico que poderá facilitar estudos explorativos ou uma fase preparatória para a construção de um tesouro, porque permite conhecer o núcleo conceitual através dessas duas teorias.

As **metodologias para tratamento de léxico** estiveram presentes nos seguintes artigos: “*Semi-automatic extraction of multiword terms from domain-specific corpora*”; “*Impact for whom? Mapping the users of public research with lexicon-based text mining*”; e “*Análisis del léxico de las emociones: una aproximación desde la disponibilidad léxica y la teoría de los grafos léxicos*”. O primeiro tratou de uma metodologia para gerir e analisar os léxicos terminológicos, ou seja, aquelas palavras e vocábulos extraídos e estabelecidos de modo a integrar plataformas e documentos de diferentes áreas e especialidades.

Pajić e colaboradores (2018) destacaram a integração de procedimentos voltados para a seleção do corpus de uma língua (no caso o idioma sérvio) por meio de análise e busca de padrões, utilizando um sistema de processamento de corpus baseado em uma tecnologia orientada e automática - o software Unitex. A abordagem e a metodologia utilizadas podem ser usadas para outras línguas, necessitando adaptar o tratamento para a língua escolhida e suas estruturas sintáticas, pois estas diferem de acordo com o léxico e as regras linguísticas. Salienta-se que os autores estiveram seguindo etapas de análise, seleção e tratamento do léxico a partir de dicionários e o emprego de PNL, dentre os quais a técnica de lematização que busca termos mais precisos durante a aplicação da vetorização.

O segundo artigo, apontou uma **metodologia baseada em mineração de textos aplicada ao léxico** diante da combinação de ferramentas e procedimentos do PNL (Bonaccorsi; Chiarello; Fantoni, 2021). Percebeu-se que a priori foram investigadas e selecionadas as unidades de análise (palavras e documentos), conforme as medidas e os índices de similaridades, dos quais permitiram elencar essas unidades em redes (mapas). Outros pontos de destaque foram: a aderência à abordagem *top-down*; e a integração de aspectos metodológicos

fixados em análise da literatura e nos dados estatísticos de acordo com informações fornecidas pelos usuários. A experiência é mais complexa, pois o estudo integrou diferentes aportes linguísticos e computacionais.

Na terceira publicação metodológica exclusiva para léxicos, os autores, apesar de demonstrarem um projeto piloto, esboçaram um processo que incluíram os métodos de análises de disponibilidade léxica e a Teoria dos Grafos Lexicais. A metodologia esteve resumida em alguns passos, relatados por Correa, Lagos e Grabole (2020):

- 1) Disponibilidade lexical – com base nos critérios de edição de dados, permitindo a unificação dos diferentes léxicos (eliminação de termos repetidos, correção e unificação ortográfica, unificação de formas completas e abreviações, unificação de verbetes compostos por diversos elementos e suas variantes abreviadas);
- 2) Análise utilizando o programa *Dispogen* – possibilitou a apresentação de dados relativos ao número total de palavras, ao número de palavras diferentes, ao número médio de palavras e ao índice de coesão para cada unidade lexical;
- 3) Aplicação da Teoria dos grafos lexicais por meio programa *Dispografo* – observação das conexões existentes entre cada léxico “emocional”, dentre os quais ao verificar hiperonímia, sinonímia ou causa e efeito dos significados, e a obtenção das relações e redes semânticas conforme os nós-arestas e a formação de *clusters*.

Os procedimentos se tornaram uma opção para medir a relação entre termos em virtude dessa teoria analisar as conexões entre as unidades. Além disso, dedicou uma etapa para reconhecer aspectos semânticos relativos aos sentimentos do público estudado. Diante desse trabalho e de alguns anteriormente citados, a integração entre etapas e instrumentos metodológicos com usuários a partir dos léxicos e das terminologias foram estabelecidas. Isso demonstrou ser uma tendência promissora aplicada em domínios científicos e em outros âmbitos sociais não especializados. A forma como essa tendência permitiu coletar dados da experiência e do uso de palavras e termos, possibilita gerar um material conciso em relação à precisão de como essas unidades são comunicadas utilizadas e como elas impactam nos requisitos funcionários de sistemas, sites, repositórios ou bibliotecas.

A publicação “*Sentiment classification: a lexical similarity based approach for extracting subjectivity in documents*” esteve pautada em uma **metodologia simples e estatística** com métodos já mencionados, anteriormente, embora ela buscou a classificação de sentimentos

para monitorar um conjunto de unidades lexicais. Em outros termos, Sarvabhotla, Pingali, Varma (2011) apontaram essa metodologia para extrair a subjetividade (método de estimar a subjetividade foi chamado de RSUMM) de cada frase de um determinado documento, utilizando técnicas semelhantes ao modelo de espaço vetorial em Recuperação da Informação. Essa metodologia é bem próxima da noção de um modelo de similaridade lexical entre vetores e as frequências (subjetivas) medidas entre termos.

Os padrões foram os últimos tópicos analisados. A pesquisa tentou verificar todos os documentos de todas as temáticas, embora os padrões foram identificados (Quadro 13) somente na literatura em tesauros.

Quadro 13 – Padrões e modelos mais destacados na literatura de construção de tesauros

Padrões	Autores
ISO e ANSI/NISO (histórico)	Krooks e Lancaster – 1993; Sánchez-Cuadrado, Colmenero-Ruiz e Moreira - 2012; Clarke e Zeng – 2012;
ISO 2788	Caminotti e Martínez – 2006; Sánchez-Cuadrado, Colmenero-Ruiz e Moreira – 2012;
BS 8723	Shuter – 2008; Calvitti e Viti – 2009; Motta e Tiberi – 2009;
ANSI/NISO Z39.19-2005	Calvitti e Viti – 2009; Hariri, Kazerani e Moghaddam-Kiya – 2013; Weiss e Bräscher;
ISO 25964	Clarke – 2008; Calvitti e Viti – 2009; Clarke e Zeng – 2012; Will – 2012; Sánchez-Cuadrado, Colmenero-Ruiz e Moreira – 2012; Alexiev, Isaac e Lindenthal – 2016; Weiss e Bräscher – 2017; Martínez-González e Alvite-Díez – 2020; Ferreira e Maculan – 2020
SKOS (Simple Knowledge Organization System)	Clarke e Zeng – 2012; Will – 2012; Sánchez-Cuadrado, Colmenero-Ruiz e Moreira – 2012;

Fonte: Elaborado pela autora.

Diante do quadro geral de padrões, Krooks e Lancaster (1993) resgataram a evolução das diretrizes e princípios para a construção de tesauros de recuperação entre os anos 1959 a 1993, concluindo que a maioria dos problemas básicos da construção de tesauros estava identificada e relacionada aos avanços técnicos, dos quais várias questões foram resolvidas em iniciativas e orientações a partir de 1967. As observações desse período estiveram encaminhadas na comunidade profissional e acadêmica para facilitar a elaboração de tesauros e vocábulos multilíngues.

Clarke (2008), Clarke e Zeng (2012) e Sánchez-Cuadrado, Colmenero-Ruiz e Moreira (2012) apontaram aspectos evolutivos em relação aos padrões internacionais e

nacionais, enfatizando a recente ISO 25964 e os esforços em direção aos sistemas e vocabulários interoperáveis, além do modelo vigente, SKOS (Simple Knowledge Organization System), e as ações do Consórcio World Wide Web (W3C).

Em relação às anteriores, os padrões BS 8723 (2005 e 2008), ANSI/NISO Z39.19-2005 e ISO 25964 introduziram outros tipos de vocabulários além dos tesouros (listas, anéis de sinônimos, taxonomias, listas de autoridades, ontologias e tabelas de classificação), buscando adaptação para o compartilhamento e aproveitamento de dados não somente entre sistemas, mas vislumbrando a disposição da indexação e recuperação da informação nos ambientes digitais. Shuter (2008) apontou aspectos gerais dos padrões 8723, de 2007 e 2008, em relação à cobertura de assuntos e aos tipos de vocabulário, além de salientar que a ISO 999 precisaria se adequar a indexação, em especial a indexação não linear.

Calvitti e Viti (2009) demonstraram que os novos padrões atendem as modernas possibilidades de pesquisa oferecidas pela Web, relatando que as normas inglesas (BS 8723 de 2005 e 2008) e as norte-americanas (ISO 25964 e ANSI/NISO Z39.19-2005) oferecem modelos e métodos de mapeamento de vocabulários estruturados com finalidade de possibilitar uma integração entre diferentes ferramentas através da interoperabilidade. As autoras vão ao encontro de Motta e Tiberi (2009) ao analisar o padrão britânico 8723 e o tesouro italiano *Thesaurus del Nuovo Soggettarlo*, e enfatizaram, além das atualizações citadas anteriormente, a adaptação da prática em vocabulário multilíngues e os critérios de análise facetada.

Martínez-González e Alvite-Díez (2020) deixaram claro que o padrão ISO 25964 evoluiu de modo que os conceitos nos tesouros são bem alinhados com a abordagem da Web Semântica, se comparado com os tesouros baseados na ISO 2788.

Hariri, Kazerani e Moghaddam-Kiya (2013) utilizaram as orientações do padrão ANSI / NISO 39.19-2005 para a construção e manutenção de tesouros persas, observando que, durante a elaboração de formulários e as composições das relações semânticas, as orientações não deram a devida atenção às dimensões de gerenciamento e exibição. Weinberg, em 2009, teceu considerações em relação ao padrão Z39.19, deixando claro observações e inconsistências quanto ao NISO de 2005 para vocabulários controlados.

Caminotti e Martínez (2006) concluíram, a partir do desenvolvimento de tesouro e taxonomia do domínio do Futebol, que as normas ISO 2788 e a norma nacional IRAM 32057 (Instituto Argentino de Normalización, 1983) foram úteis para melhorar a precisão na recuperação da informação, sem alterar a terminologia típica do ambiente futebolístico. A norma argentina e o padrão ISO orientam as formas gramaticais e o controle dos sinônimos, homonímia e polissemia. Assim como conforme a leitura de Gomes e Campos, em 2004, as

soluções para os tesouros e da normalização terminológica estão inseridas na ISO 2788, das quais tomam como subsídios à formação gramatical e às formalidades em que as terminologias estão submetidas no processo de comunicação e no contexto universal da ciência.

Analisando aspectos mais recentes, Will (2012) afirmou que a ISO 25964 facilita a apresentação de um sistema de organização do conhecimento em SKOS (Simple Knowledge Organization System) compatível com máquinas, fornecendo estratégias para além das simples relações hierárquicas e conceitos relacionados, pois evidencia o uso de tesouros para recuperação de informações, incluindo equivalência composta e rótulos de nós.

Alexiev, Isaac e Lindenthal (2016) retomaram que as relações hierárquicas na ISO 25974 são utilizadas nos sistemas de organização por meio dos indicadores: genéricas mais amplas (BTG), partitivas mais amplas (BTP) e instanciais mais amplas (BTI), assim esses tipos de relacionamentos especializados não podem ser combinados arbitrariamente para estabelecer declarações que tenham a mesma precisão semântica e levam a casos em que a inferência dos relacionamentos transitivos mais amplos podem causar imprecisão. Ou seja, a questão citada pelos autores se remete, também, à composição de relações hierárquicas, das quais devem ser mantidas para não ocasionar a perda e a ineficiência da expansão dos termos e das propriedades semânticas durante a realização da pesquisa ou escolha de um determinado termo.

Em 2017, Weiss e Bräscher foi possível reconhecer que às recomendações ANSI/NISO Z39.19(2005) e ISO 25964-1 corresponderam à abordagem da semântica pragmática conforme o entendimento e estruturação das relações semânticas. Embora não foi viável identificar se de fato esses padrões interferem ou inviabilizam relações semânticas em contexto de uso de léxicos derivados de grupos sociais ou de influências étnico-raciais. Ainda nas relações entre termos, Ferreira e Maculan, em 2020, expuseram a ISO 25.964-1 (2011) em comparação aos critérios para elaboração de relações semânticas hierárquicas, associativas e de equivalência em tesouros.

A partir desses estudos e dos documentos dos padrões vigentes ISO 25.964, de 2011 e de 2013, verificou-se que essas orientações se adequam ao tratamento e ao estabelecimento dos termos de especialidades e de exemplos gerais dos léxicos das línguas mais gerais, principalmente o inglês e o espanhol, embora os relatos e observações foram insuficientes para afirmar o êxito e a eficiência em léxicos que sofrem com variantes e os fatores linguísticos regionais e dialetos, dentre eles as influências étnicas.

6.4 Síntese

Ao verificar as convergências, as aproximações e os afastamentos relatados em cada categoria, as subcategorias (conceitos, teorias, abordagens, métodos etc.) representaram possibilidades a serem analisadas, em termos teóricos e práticos de construção de tesauros, juntamente, com o aporte da Análise de Domínio. A Figura 17 assimilou esses elementos.

Figura 17 – Síntese dos aspectos conceituais, teóricos e metodológicos a partir da aplicabilidade da Análise de Domínio e da leitura de tópicos étnico-raciais e culturais



Fonte: Elaborado pela autora.

Essas subcategorais e suas respectivas funções e aplicabilidades são possíveis elementos para compor um modelo de estudo ou modelo prático de um determinado tesauro de acordo com o domínio, a comunidade e o cenário. Entretanto, há pontos divergentes, relatados

anteriormente, dos quais não viabilizam determinadas combinações teórico-metodológicas. Como exemplos, as investigações ambientadas em espaços onde existe a presença de variantes terminológicas e léxicas (moldadas por aspectos culturais) que não condizem com abordagem ou o paradigma cognitivo; ou os métodos de coleta e análises oriundos da Teoria Behaviorista na Linguística porque não viabilizam o reconhecimento de palavras e vocábulos de um determinado grupo social conforme os fatos históricos e culturais.

6.5 Verificação das Hipóteses

Inicialmente, esta pesquisa apontou duas hipóteses essenciais para alcançar a resposta da questão de pesquisa. A seguir, elas são verificadas de acordo com a análise e discussão dos dados.

a) **Hipótese 1** - A fundamentação e os aspectos metodológicos aplicados à elaboração de Tesouros em conjunto com o aporte teórico-conceitual e metodológico da Análise de Domínio se complementam e dialogam com a prática representacional e organizacional em domínios científicos e não científicos.

As publicações caracterizadas pela Análise de Domínio estiveram presentes nos estudos dos sistemas de organização do conhecimento e, em específico, os estudos teóricos e práticos dos tesouros. Uma parcela expressiva de autores utilizou a AD enquanto paradigma social da CI e enquanto abordagem teórica ou metodológica, representando 65 dos 139 trabalhos analisados (46,76 %). Desses títulos, 35 confirmaram e ou apontaram a AD como aporte teórico e metodológicos direcionados à construção e à manutenção de tesouros. Além desses, mais 6 títulos se dedicaram, exclusivamente, ao desenvolvimento de sistemas de organização do conhecimento e 3 trabalhos se destinaram à representação temática da informação (indexação) e à representação descritiva da informação (catalogação).

Aferiu-se que a Análise de Domínio, conforme seus postulados e de acordo com a noção de paradigma de Kuhn (2005), está ainda em um processo evolutivo e de revolução científica em diversas frentes epistemológicas da CI. Entende-se que a AD está consolidada na CI e na Biblioteconomia quanto aos estudos da Organização do Conhecimento, mas carece de uma aderência em outros temas e subcampos. Isso tendo em vista a afirmativa de Hjørland e Albrechtsen (1995), Hjørland (2002a, 2004b, 2017) Hjørland e Hartel (2003) dessa perspectiva ser um paradigma social e uma abordagem na CI.

Segundo a análise e discussão dos dados a partir dos quadros das seções anteriores (7, 8, 9, 10, 11 e 12), constatou-se que os aspectos metodológicos e de fundamentação dos

tesauros dialogam e são influenciados pelas teorias, abordagens e métodos destacados nos estudos das análises dos domínios. Alguns autores estiveram demonstrando a aplicabilidade do aporte teórico, conceitual e metodológico da AD para a consolidação dos tesauros enquanto um vocabulário estruturado, visualizada desde:

- a) O resgate das Teorias do Conceito e da Classificação Facetada presentes nos princípios e pressupostos para estruturar, conceitualmente, os termos e suas relações;
- b) O descobrimento de domínios segundo a análise de unidades conceituais e o suporte no tratamento terminológico;
- c) A compreensão das comunidades discursivas (produtores e consumidores de informação e de conhecimento) e sua representatividade nas dimensões social e cognitiva de uma temática, disciplina ou interesse – ambos elementos candidatos ao serem domínio segundo Mai (2008), Smiraglia (2011, 2015a) e Hjørland (2017).

Acerca dos domínios científicos e não científicos, retoma-se a noção de disciplina e *habitus* científico de Bourdieu (2004a, 2004b) e a institucionalização social e cognitiva destacadas por Whitley (1974) em virtude do estabelecimento dos processos e do desenvolvimento científico. Nesse sentido, um domínio científico estaria esclarecido naquilo que para Bourdieu (2004a) se concentra no campo científico estabelecido segundo a distribuição do capital científico e cuja constituição de uma disciplina se dá a partir da sua autonomia. Desse modo, segundo o mesmo autor (Bourdieu, 2004b) uma disciplina é uma delimitação de um campo científico e ocupa um *nomos* particular, um princípio de visão e divisão, um princípio de construção da realidade objetiva e irreduzível ao de outras disciplinas, estando em uma estrutura hierarquizada e institucionalizada. Diferentemente dessa noção, um domínio não científico emerge de outras variações estruturais decorrentes de outros padrões de hierarquização e de socialização do conhecimento, surgindo, predominantemente, em contextos coletivos e não universais.

Em Whitley (1974), os domínios podem ser percebidos de acordo com o compartilhamento de seus construtos científicos (pensamentos, modelos, teorias, propostas) em duas dimensões (cognitiva e social). Nas disciplinas científicas essas dimensões possuem sistemas hierárquicos baseados em ganhos e compartilhamentos de seus produtos intelectuais e científicos, dos quais todos os membros das comunidades precisam seguir protocolos de

formação, desempenho e busca por ascensão conforme a aquisição e a troca de conhecimento. As disciplinas não científicas são semelhantes, contudo, sua sistemática é pautada em outros aspectos, dentre eles os culturais que podem variar de acordo com o lugar e o grupo.

Assim, para domínios relativos ao conhecimento não científico (popular, religioso), essas duas proposições anteriores (Bourdieu, 2004a; Whitley, 1974) não se aplicam em sua completude, principalmente, por serem caracterizados e fundamentados em outras estruturas. Nessa perspectiva, os domínios mencionados e eleitos pelos autores foram distinguidos no Quadro 14.

Quadro 14 - Domínios científicos e domínios não científicos

Domínios científicos
Cultura na Organização do Conhecimento - Rodrigues, Anderson Luiz Cardoso, 2012
Estudos de gêneros documentais (Arquivologia) - Hernandez, Bianca Ferreira; Vital, Luciane Paula, 2022
Faculdade de Ciências Naturais e Museu da Universidade Nacional de La Plata - Miguel, Sandra; Moya-Anegón, Félix; Herrero-Solana, Víctor, 2008
Informação digital (Bibliotecas digitais sob a perspectiva da Biblioteconomia e da Ciência da Informação) - Lee, Jae Yun; Kim, Heejung; Kim, Pan Jun, 2010
Domínios não científicos
Contra Dance (Dança Country) - Coladangelo, L. P., 2020
Indústria Farmacêutica (Atividade comercial da área da Saúde) - Nielsen, Marianne Lykke, 2001
Grafite (Grafite, Grafito) - Graf, Ann M., 2022
Culinária gourmet - Hartel, Jenna, 2010
Artefacts Canada Humanities - Menard, Elaine; Mas, Sabine; Alberts, Inge, 2010
Cultura Amazônica - Rodrigues, Anderson Luiz Cardoso, 2012

Fonte: Elaborado pela autora.

Os domínios não científicos identificados pertencem às temáticas, comunidades e ambientes informacionais distintos daqueles outros domínios oriundos da institucionalização e dos campos científicos estabelecidos em instituições acadêmicas e grupos científicos consolidados. O Grafite é um exemplo de como uma arte urbana se consolida enquanto manifestação artística, amplamente difundido por comunidades em diferentes regiões, estados e países e cujo conjunto de conhecimento e técnicas que os caracterizam são unidos e convergem para traços artísticos e estilos comuns compartilhados por seus membros. Assim, o Grafite expressa a qualidade de um domínio conforme a proposição de Hjørland (2017, p. 6, tradução nossa) de que “[...] domínios são frequentemente disciplinas acadêmicas, mas também podem ser, por exemplo, *hobbys*”, e de Smiraglia (2011, 2012, 2015a) de que esses apresentam

uma semântica social, consensos comuns, corpo de conhecimento e epistemologia. Nesse caso, a epistemologia não é reconhecida enquanto científica, mas apresenta seus alicerces em técnicas e tecnologias consolidadas em conhecimento implícitos e explícitos de coletivos e grupos urbanos.

b) **Hipótese 2** - Quanto à construção de tesouros que utilizam termos oriundos do léxico de influência étnico-racial e as investigações da Análise de Domínio na Ciência da Informação, os aportes teóricos e metodológicos se apresentam insuficientes para tratar e investigar tal realidade étnico-racial no Brasil.

O mapeamento dos dados evidenciou, desde a etapa de pré-seleção, um baixo número de pesquisas que trataram dos léxicos das línguas naturais nos tesouros, destacando os estudos voltados aos grupos socialmente minorizados, povos étnico-raciais, assim como aspectos históricos e culturais. Essa ausência e insuficiência de pesquisas na Ciência da Informação e na Biblioteconomia, também, esteve ainda mais expressiva no conjunto de títulos que abordavam a temática de construção de tesouros e AD.

Logo, comprovou-se que os aportes teóricos e metodológicos não atentam ou incorporam, de modo completo ou expressivo, às questões referentes ao léxico de influência étnico-racial e seus possíveis usos no desenvolvimento de sistemas de organização do conhecimento. Outra questão de exiguidade esteve ligada aos poucos títulos que promoveram diálogos interdisciplinares entre a Linguística, Sociolinguística, Socioterminologia e outras, como a Semântica Discursiva, Teoria Comunicativa da Terminologia e a Psicossociologia no que confere os estudos as representações sociais e das questões sociocognitivas.

Embora, uma parcela de trabalhos foi considerada imprescindível para firmar o relacionamento e as possibilidades, dentre eles propondo: o método e a abordagem etnográfica ao encontro de estudos empíricos que mapeiam e observam os léxicos e as unidades conceituais; os métodos bibliométricos para visualizar comunidades discursivas, terminologias, conceitos, frentes e escolas epistêmicas; os estudos críticos ao encontro das investigações conceituais e terminológicas acerca das epistemologias decoloniais. Esses são alguns dos exemplos de metodologias e suportes epistêmicos que a Análise de Domínio auxilia na fundamentação dos tesouros, e que podem ser úteis e viáveis para discutir as questões lexicais étnico-raciais no Brasil.

7 CONCLUSÃO

O estabelecimento dos tesouros percorreu uma evolução histórica marcada por constantes transformações. Nessa trajetória, esses vocabulários estruturados estiveram produzidos e conduzidos para fins e interesses científicos e políticos, uma vez que as principais vertentes (europeias e estadunidense) limitou seus esforços para desenvolver sistemas e ferramentas de organização da informação e do conhecimento destinadas às organizações institucionais e comerciais, como no caso das pesquisas das empresas Du Pont e da IBM e dos centros governamentais britânico e estadunidense.

Recorrendo ao objetivo de examinar aspectos teóricos, conceituais e metodológicos basilares na construção do tesouro, verificou-se que o propósito desse instrumento de recuperação, tido na literatura enquanto um sistema de organização do conhecimento, esteve mais distante das políticas e soluções ofertadas às comunidades e à sociedade geral e mais próximo das elites científicas e governamentais. Isso implica em indícios de uma fundamentação de epistemologias hegemônicas que são incorporadas ao avanço do positivismo e de suas ideias de processo e universalização, além da presença dos métodos e abordagens cognitivistas. Ou seja, deixando de lado outras perspectivas sociocognitivas, epistemologias decoloniais e saberes locais particularizados em cada língua, etnia e cultura.

A Teoria da Terminologia Geral e a completude com que ela esteve difundida nos primeiros padrões nacionais e internacionais de construção de terminologias e vocabulários, condicionaram por muito tempo e, ainda são influentes, nessa fundamentação tradicional dos tesouros. A TGT, também, denota alguns pontos elucidados na Teoria do Conceito (a lógica formal e o princípio de univocidade). A teorização dos conceitos conseguiu contribuir com as investigações dos fenômenos e objetos científicos enquanto construção conceitual. Embora, os postulados da lógica podem ser criticados por algumas abordagens e vertentes.

Por outro lado, a Teoria da Classificação Facetada ampliou o plano teórico e prático discutido na literatura, principalmente, por reconhecer que o conhecimento é uma espiral, um universo, assim como o universo das ideias apresenta a elucidação das experiências e reverbera na atribuição dos conceitos. O que para alguns estudos de esquemas e de vocabulários, a herança facetada é algo complexo e de difícil aplicabilidade.

Atendendo ao objetivo de investigar a Análise de Domínio na CI, essa esteve empregada em alguns estudos, não somente no desenvolvimento prático de alguns componentes ou elementos que contemplam o desenvolvimento e manutenção dos tesouros, mas reiterou investigações que perpassaram desde o planejamento, ao descobrimento da temática (aqui

defendida como domínio) até a compreensão das comunidades (discursivas e epistêmicas), sejam de especialistas e produtores de conhecimento, sejam de usuários que utilizam e ressignificam os termos diante do uso desses vocabulários estruturados.

Em síntese, podemos considerar o domínio como sendo uma unidade constituída por repertórios de conhecimento e informação resultantes da interação entre: atores de uma comunidade ou grupo com aproximações sociais e culturais, com uma determinada terminologia ou semântica social, estruturas e organização disciplinar para estabelecer atividades e saberes comuns. Nessa noção, estão agregadas aproximações sociais e culturais por considerá-las fatores que implicam ou não nos consensos e dissensos, além de outras interrogações voltadas para as ideologias, vertentes políticas, doutrinas e religiões.

Contudo, alguns métodos empíricos da AD e métodos automáticos e estatísticos podem ser inconsistentes durante a aplicação em determinados domínios e realidades, ou serem insuficientes por contemplar somente algumas etapas ou aspectos de determinados estudos e práticas. É necessário, então, verificar qual a base epistêmica do domínio pode ser mais adequada a determinados métodos e abordagens do que outros aportes metodológicos.

Quanto ao objetivo de investigar os estudos e as frentes epistemológicas que discutem e contribuem com os usos e o entendimento dos léxicos influenciados pelos aspectos étnicos, foram apontados os aspectos teóricos da etnicidade e dos conceitos basilares das questões étnico-raciais, acrescidos dos estudos da língua e do léxico de acordo com a Linguística, a Lexicologia, a Lexicografia e a Terminologia. Demonstrando que, assim como a língua, as relações de poder estão enraizadas nas estruturas do sistema moderno e no modelo do colonialismo ao impor a classificação social e os mecanismos de violência e de dominação, ambos com ações e comportamentos formalizados através do racismo.

A introdução das questões étnicas e dos léxicos das línguas indígenas e afro-brasileiras, ainda que de forma breve, possibilitou compreender a urgência de resgatar e dar importância aos elementos e resquícios que representam os grupos e as comunidades denominadas nas literaturas étnico-raciais. Isso está ligado a um dos propósitos desta pesquisa: o de contribuir com os diálogos e as pesquisas em organização e representação (do conhecimento e da informação) voltados à promoção da preservação das memórias e à introdução dos elementos culturais e dos saberes desses povos diante de um cenário político e social enfraquecido pelo silenciamento e pela extinção de indígenas e demais etnias afrodescendentes no Brasil.

Os conceitos de etnoconhecimento, colonialidade do poder, racismo, etnia, decolonialidade, classificação social e das questões de diversidade de gênero e identidades,

requer um aprofundamento em relação à fundamentação teórica e prática dos tesouros, pois os padrões e princípios vigentes não abordam e atendem aos seus aspectos e suas especificidades, além das complexidades frente a sua organização e sua representação, seja social, cultural, linguística, religiosa ou política.

Diante do objetivo geral, as reflexões geradas indicam que a Análise de Domínio parte das preocupações gerais da CI e encontram-se com preocupações típicas daquelas ressaltadas pelos profissionais que elaboram e administram os tesouros, como a dinâmica de evolução de unidades conceituais quando as especialidades e as temáticas, mas não se encontram, de forma consistente, nas discussões e nos diálogos da literatura dos léxicos e das influências das línguas étnicas. Ou seja, há poucas pesquisas na CI e na Biblioteconomia nesse sentido. Outros assuntos podem ser explorados e recorrer às outras teorias terminológicas (Teoria Comunicativa da Terminologia e o enfoque cultural na Terminologia) e à Socioterminologia.

A partir dessas constatações, a tese levantada inicialmente atestou que a Análise de Domínio aporta elementos necessários para a reestruturação teórico e metodológica, já que integra a dimensões (ontológica, epistemológica e sociológica) destinadas à investigação de domínios científicos e não científicos, como os domínios ou assuntos acerca do léxico do português brasileiro estabelecido pelas heranças linguísticas dos povos indígenas e afro-brasileiros. Esses domínios são candidatos a serem objetos de estudo e foco das construções de tesouros por representarem e permitirem o reconhecimento e o tratamento das unidades conceituais frutos do compartilhamento de saberes e de uma semântica social compartilhada por uma comunidade.

A pesquisa percorreu com algumas limitações. A primeira em virtude das consequências causadas pela pandemia da COVID-19 em diferentes esferas, das quais influenciaram a configuração metodológica deste trabalho. A segunda esteve relacionada ao acesso aos documentos impressos que não puderam ser lidos e incorporados aos capítulos do referencial teórico. Diante dessas dificuldades, o estudo poderá ser explorado em alguns pontos introdutórios, aqui esboçados, ao surgir outras particularidades. A continuidade desse trabalho está em futuras pesquisas acerca de aplicar aos aspectos metodológicos e teóricos para elaboração de um tesouro temático de uma etnia indígena do estado do Ceará, além de explorar outras lacunas não contempladas nesta tese. Uma das lacunas se refere ao estudo para investigar os padrões internacionais vigentes com o foco exclusivo na incorporação de termos indígenas brasileiros, propondo melhorias em tesouros bilíngues e multilíngues acerca da cultura brasileira.

REFERÊNCIAS

- AGUILERA, Vanderci De Andrade; SILVA, Hélien Cristina da. Tecendo a história do léxico da língua portuguesa no Brasil com os fios das telas dos atlas linguísticos. **Working Papers em Linguística**, v. 23, n. 1, p. 168-193, 2022.
- AITCHISON, Jean; CLARKE, Stella G. Dextre. The Thesaurus: A Historical Viewpoint, with a Look to the Future. **Cataloging & Classification Quarterly**, v. 37, n. 3-4, p. 5-21, 2004. DOI: http://doi.org.10.1300/J104v37n03_02. Acesso em: 2 fev. 2021.
- AITCHISON, Jean; GILCHRIST, Alan. **Manual para construção de tesauros**. Rio de Janeiro: Brasilart, 1979.
- ALBRECHTSEN, H. Subject analysis and indexing: from automated indexing to domain analysis. **The Indexer**, v.18, n.4, p.219-224, out., 1993.
- ALBRECHTSEN, Hanne. This is not domain analysis. **Knowledge Organization**, v. 42, n. 8, p. 557-561, 2015.
- ALBUQUERQUE, Ana Cristina de. Por um reflexão sobre a organização e representação de conceitos decoloniais na América Latina: O pensamento de Aníbal Quijano à luz da Análise de Domínio. **Encontros Bibli**, v. 28, p. e92960, 2023.
- ALEXIEV, Vladimir; ISAAC, Antoine; LINDENTHAL, Jutta. On the composition of ISO 25964 hierarchical relations (BTG, BTP, BTI). **International Journal on Digital Libraries**, v. 17, p. 39-48, 2016.
- ALKIMIM, Tania; PETTER, Margarida Maria Taddoni. Palavras da África no Brasil de ontem e de hoje. In: PETTER, Margarida; FIORIN, José Luiz. **África no Brasil: a formação da língua portuguesa**. São Paulo: Contexto, 2014, p. 145-178.
- ALMEIDA, Daniela Pereira dos Reis de et al. Paradigmas contemporâneos da Ciência da Informação: a recuperação da informação como ponto focal. **Revista Eletrônica Informação e Cognição** (Cessada), v. 6, n.1, p.16-27, 2007.
- ALMEIDA, Maurício Barcellos; EMYGDIO, Jeanne Louize. Uma investigação teórica sobre relações semânticas partitivas e sua aplicação em sistemas de organização do conhecimento. **Informação & Informação**, v. 24, n. 2, p. 31-37, 2019.
- ALVARENGA, Lídia; DIAS, Célio da Consolação. Análise de Domínio e Gestão Arquivística. **DataGramaZero**, v. 13, n. 1, 2012.
- AMERICAN CHEMICAL SOCIETY. **Chemical Abstracts Service**: National Historic Chemical Landmark, 2021. Disponível em: <https://www.acs.org/content/acs/en/education/whatischemistry/landmarks/cas.html>. Acesso em: 17 dez. 2021.
- AMERICAN LIBRARY ASSOCIATION. **History**. 2021. Disponível em: <https://www.ala.org/aboutala/node/230/>. Acesso em: 2 dez. 2021.
- AMERICAN NATIONAL STANDARDS INSTITUTE. National Information Standards Organization. **ANSI/NISO Z39.19-2005 (R2010)**: guidelines for the construction, format and management of monolingual controlled vocabularies. Bethesda: NISO Press, 2010.
- AMORIM, Igor Soares; CAFÉ, Lígia Maria. Agenciamento e análise de domínio: um encontro possível. **Informação & Sociedade**, v. 27, n. 2, 2017. Disponível em:

https://www.brapci.inf.br/_repositorio/2017/10/pdf_a2d6319084_0000027007.pdf. Acesso em: 3 jul. 2023.

AMORIM, Igor Soares; VIANNA, William Barbosa; MEDEIROS, Marisa Brasher Basilio. Aspectos epistemológicos em estudos sobre Análise de Domínio sob as lentes do racionalismo e do empirismo. **Logeion: Filosofia da Informação**, v. 5, n. 1, p. 35-47, 2018.

BARBOSA, Maria Aparecida. Pesquisas em etnoterminologia: o repertório terminológico e cultural em regiões amazônicas. In: XIII Congresso Nacional de Linguística e Filologia, 2009, RJ. **Livro de Resumos Programação do XIII Congresso Nacional de Linguística e Filologia**. Rio de Janeiro: CIEFIL, 2009. v. 1, p. 139-139.

BARDIN, Laurence. **Análise de conteúdo**. 1 ed. revista e ampliada. São Paulo: Edições 70, 2011.

BARITÉ, Mario et al. Garantia literária: elementos para uma revisão crítica após um século. **TransInformação**, v. 22, p. 123-138, 2010.

BARROS, Thiago Henrique Bragato. Michel Pêcheux's Discourse Analysis: An Approach to Domain Analyses. **Knowledge Organization**, v. 50, n. 1, p. 3-10, 2023.

BARROSO, Danielle *et al.* **Epistemologias negras**: relações raciais na biblioteconomia. Florianópolis: Rocha Gráfica e Editora, 2019.

BARTH, Fredrik. **O guru, o iniciador e outras variações antropológicas**. Rio de Janeiro: Contra Capa Livraria, 2000.

BEGHTOL, Clare. Domain analysis, literary warrant, and consensus: the case of fiction studies. **Journal of the American Society for Information Science**, v. 46, n. 1, p. 30-44, 1995.

BENTES PINTO, Virginia; CAMPOS, Henry de Holanda; ARAGÃO, Maria do Socorro; RIBEIRO, Cristina da Costa. O léxico das comunidades indígenas do Ceará na designação de doenças: reflexões para a construção de vocabulário controlado. **Perspectivas em Ciência da Informação**, v. 25, n. 2, p.171-193, jun. 2020. DOI: <https://doi.org/10.1590/1981-5344/381> Acesso em: 30 jul. 2021

BERGER, Petter L.; LUCKMANN, Thomas. **A construção social da realidade**. Petrópolis: Vozes, 2014.

BERNARDINO-COSTA, Joaze; MALDONADO-TORRES, Nelson; GROSGOUEL, Ramón. **Decolonialidade e pensamento afrodiaspórico**. Belo Horizonte: Autêntica, 2018.

BERNIER, Charles L. Indexing and thesauri. **Special Libraries**, v. 59, n. 2, p. 98-103, 1968.

BIDERMAN, Maria Tereza Camargo. As ciências do léxico. In: OLIVEIRA, Ana Maria Pinto Pires; ISQUERDO, Aparecida Negri. (Orgs.). **As ciências do léxico**: lexicologia, lexicografia, terminologia – v. 1. Campo Grande: Ed. UFMS, 2001.

BIDERMAN, Maria Tereza. O léxico testemunha de uma cultura. In: HOINKES, Ulrich; DIETRICH, Wolf(Orgs.). **Kaleidoskop der Lexikalischen Semantik**. Gunter Narr Verlag Tübingen, 1997, p. 308-399.

BONACCORSI, Andrea; CHIARELLO, Filippo; FANTONI, Gualtiero. Impact for whom? Mapping the users of public research with lexicon-based text mining. **Scientometrics**, v. 126, p. 1745-1774, 2021.

- BONVINI, Emilio. Línguas africanas e português falado no Brasil. In: PETTER, Margarida; FIORIN, José Luiz. **África no Brasil: a formação da língua portuguesa**. São Paulo: Contexto, 2014, p. 15-62.
- BORBA, Francisco S. Propriedades sintáticas, semânticas e pragmáticas do léxico. *Revista (Con) textos linguísticos*, v. 1, n. 1, p. 55 – 68, 2007.
- BORKO, H. Information Science: What is it? *American Documentation*, v.19, n.1, p.3-5, Jan. 1968.
- BOURDIEU, Pierre. **Os usos sociais da ciência: por uma sociologia clínica do campo científico**. São Paulo, SP: Editora UNESP, 2004a.
- BOURDIEU, Pierre. **Para uma sociologia da ciência**. Lisboa: Ed. 70, 2004b.
- BRAGGIO, Silvia Lucia Bigonjal. Variedade dialetal do português em contato com uma língua indígena. In: **Estudios de lengua y cultura amerindias II: lenguas, literaturas y medios: actas de las IV Jornadas internacionales de lenguas y cultura amerindias**, Valencia 17-20 de noviembre de 1997. Universitat de València, 1998. p. 169-179.
- BRÄSCHER, Marisa. Semantic relations in knowledge organization systems. *Knowledge Organization*, v. 41, n. 2, p. 175-180, 2014.
- BRÄSCHER, Marisa; CAFÉ, Lígia Maria Arruda. Organização da informação ou organização do conhecimento? In: **Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB)**, 9., Anais eletrônicos..., 2008, São Paulo. Disponível em: <https://www.brapci.inf.br/index.php/res/v/176535>. Acesso em: 4 jun. 2018.
- BROUGHTON, Vanda. Facet analytical theory as a basis for a knowledge organization tool in a subject portal. In: **Seventh International ISKO Conference, Granada, Spain**, p. 10-13, jul. 2002. Disponível em: <https://www.ucl.ac.uk/fatks/paper2.htm>. Acesso em: 2 fev. 2022.
- BROUGHTON, Vanda. **Essential thesaurus construction**. Facet publishing, 2006a.
- BROUGHTON, Vanda. The need for a faceted classification as the basis of all methods of information retrieval. *Aslib proceedings*, v. 58, n. 1/2, p. 49-72, 2006b. DOI: <https://doi.org/10.1108/00012530610648671>.
- BUFREM, Leilah Santiago. Complementaridade qualitativo-quantitativa na pesquisa em informação. *Transinformação*, v. 13, n. 1, p. 49-55, 2001. DOI: 10.1590/S0103-37862001000100004. Acesso em: 18 nov. 2021.
- CABRÉ, María Teresa. **La terminología: representación y comunicación**. Barcelona: Institut Universitari de Linguística aplicada, Universitat Pompeu Fabra, 2005.
- CABRÉ, Maria Teresa; AZEVEDO, Diego Napoleão Viana. Teorias da terminologia: descrição, prescrição e explicação. *Cadernos de Tradução*, v. 39, n.3, p. 507-558, 2019. DOI: 10.5007/2175-7968.2019v39n3p507.
- CALDAS, Raimunda Benedita Cristina; BARROS, Silvia Helena Benchimol; SILVA, Rabita Fernandes de. Aspectos interculturais do léxico traduzido entre línguas e culturas na Amazônia. In: LABORDE, Elga Pérez; UNTERNBAUMEN, Enrique Huelva; NAVES, Rozana Reigota (Orgs.) **Interculturalidade e patrimônio em contextos latino-americanos**. Campinas: Pontes, 2016, p. 121-138.
- CALVITTI, Tiziana; VITI, Elisabetta. Da ISO 2788 ai nuovi standard per la costruzione e l'interoperabilità dei vocabolari controllati: un'analisi comparativa. *Bollettino AIB (1992-2011)*, v. 49, n. 3, p. 307-324, 2009.

- CAMINOTTI, María Laura; MARTÍNEZ, Ana María. Fútbol, tesauros y taxonomías Web: desafíos del control de vocabulario. **Información, cultura y sociedad**, n. 14, p. 73-81, 2006.
- CAMPOS, Luiz Fernando de Barros; VENÂNCIO, Ludmila Salomão. Perspectivas em (in) formação: tendências e tensões entre abordagens físicas, cognitivistas e emergentes. **Transinformação**, v. 19, p. 107-118, 2007.
- CAMPOS, Maria Luiza Almeida. **Linguagem documentária**: teorias que fundamentam sua elaboração. Niterói: EdUFF, 2001.
- CAMPOS, Maria Luiza de Almeida; GOMES, Hagar Espanha. Princípios de organização e representação do conhecimento na construção de hiperdocumentos. **DataGramaZero**, v. 6, n. 6, p. 2005.
- CAMPOS, Maria Luiza Almeida; GOMES, Hagar Espanha. Metodologia de elaboração de tesauro conceitual: a categorização como princípio norteador. **Perspectivas em Ciência da Informação**, v. 11, n. 3, 2006. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/32349>. Acesso em: 13 abr. 2017.
- CAPURRO, Rafael; HJØRLAND, Birger. O conceito de informação. **Perspectivas em ciência da informação**, v. 12, p. 148-207, 2007. Disponível em: <https://www.scielo.br/j/pci/a/j7936SHkZJkpHGH5ZNYQXnC/abstract/?lang=pt>. Acesso em: 23 abr. 2011.
- CARDONA, Natalia Duque; SILVA, Franciéle Carneiro Garcês. **Epistemologias latino-americanas na Biblioteconomia e Ciência da informação**: Contribuições da Colômbia e do Brasil. Florianópolis: Rocha Gráfica e Editora, 2020.
- CARLAN, Eliana; BRÄSCHER, Marisa. Sistemas de Organização do Conhecimento na visão da Ciência da Informação. **Revista Ibero-Americana de Ciência da Informação**, v. 4, n. 2, p. 53-73, 2011.
- CAVATI SOBRINHO, Heliomar. **A representação documentária do domínio da economia**: análise de estruturas de representação em linguagens documentárias e documentos específicos de economia. 2014. 148 f. Tese (Doutorado) - Universidade Estadual Paulista, Programa de Pós-graduação em Ciência da Informação, Marília, 2014. Disponível em: https://repositorio.ufc.br/handle/riufc/19139?locale=pt_BR Acesso em: 12 mar. 2016.
- CERVANTES, Brígida Maria Nogueira. **A construção de tesauros com a integração de procedimentos terminográficos**. 2009. 209f. Tese (Doutorado) Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2009. Disponível em: <https://repositorio.unesp.br/handle/11449/103382>. Acesso em: 10 fev. 2015.
- CHEN, Sophy Shu-Jiun. Methodological Considerations for Developing Art & Architecture Thesaurus in Chinese and its Applications. **Knowledge Organization**, v. 48, n. 4, p. 267-281, 2021.
- CLARKE, Stella G. Dextre. Iso 2788+ iso 5964+ much energy= iso 25964. **Bulletin of the American Society for Information Science and Technology**, v. 35, n. 1, p. 31-33, 2008.
- CLARKE, Stella G. Dextre. Origins and Trajectory of the Long Thesaurus Debate. **Knowledge Organization**, v. 43, n. 3, p. 138-144, 2016. DOI: <http://doi.org/10.5771/0943-7444-2016-3-138>. Acesso em: 2 fev. 2021.
- CLARKE, Stella G. Dextre. The Information Retrieval Thesaurus. **Knowledge Organization**, v. 46, n. 6, 2019. DOI: 10.5771/0943-7444-2019-6-439.

- CLARKE, Stella G. Dextre; GILCHRIST, Alan; WILL, Leonard. Revision and extension of thesaurus standards. **Advances in Knowledge Organization**, v. 9, p. 215-220, 2004.
- CLARKE, Stella G. Dextre; ZENG, Marcia Lei. From ISO 2788 to ISO 25964: The evolution of thesaurus standards towards interoperability and data modelling. **Information Standards Quarterly (ISQ)**, v. 24, n. 1, 2012. Disponível em: http://eprints.rclis.org/16818/1/SP_clarke_zeng_isqv24no1.pdf. Acesso em: 3 set. 2021.
- COLADANGELO, L. P. Thesaurus and Ontology Construction for Contra Dance: Knowledge Organization of a North American Folk Dance Domain. **Knowledge Organization**, v. 47, n. 7, p. 523-542, 2020.
- COLADANGELO, L. P. Organizing controversy: Toward cultural hospitality in controlled vocabularies through semantic annotation. **Knowledge Organization**, v. 48, n. 3, p. 195-206, 2021.
- COLÍN, Marisela. Los nombres del racismo en México: una mirada léxica. **Lengua y Sociedad**, v. 22, n. 1, p. 55-80, 2023.
- COLOMBO, Stephanie; BARITÉ, Mario. Tres enfoques de bias en Organización del Conocimiento: bias neutro, bias negativo y bias positivo. **Brazilian Journal of Information Science: research trends**, v. 9, n. 2, 2015.
- CORREA, Óscar Elías Blanco; LAGOS, Pedro Salcedo; GRABOLE, Gabriela Emilce Kotz. Análisis del léxico de las emociones: una aproximación desde la disponibilidad léxica y la teoría de los grafos léxicos. **Lingüística y Literatura**, n. 78, p. 55-83, 2020.
- CORRÊA-DA-SILVA, Beatriz Carretta. O mundo a partir do léxico: reconstruindo a realidade social Mawé-Awetí-Tupí-Guaraní. **Revista Brasileira de Linguística Antropológica**, v. 5, n. 2, p. 385-400, 2013.
- COSTA, Daniela de Souza Silva. **O léxico indígena nas capitais brasileiras: um estudo geolinguístico**. 2012, 208 f. Dissertação (Mestrado em Estudos de Linguagens) - Programa de Pós-Graduação em Estudos de Linguagens, Universidade Federal de Mato Grosso do Sul, Campo Grande, 2012. Disponível em: https://alib.ufba.br/sites/alib.ufba.br/files/dissertacao_daniela_costa_o_lexico_indigena.pdf. Acesso em: 20 fev. 2023.
- DAHLBERG, Ingetrut. A referent-oriented, analytical concept theory for INTERCONCEPT. **Knowledge Organization**, v. 5, n. 3, p. 142-151, 1978a. DOI: <https://doi.org/10.5771/0943-7444-1978-3-142>.
- DAHLBERG, Ingetrut. Teoria do conceito. **Ciência da informação**, v. 7, n. 2, p. 101-107, 1978b.
- DAHLBERG, Ingetrut. Fundamentos teórico-conceituais da classificação. **Revista de Biblioteconomia de Brasília**. Brasília, v. 6, n. 1, p. 9-21, 1978c.
- DAHLBERG, Ingetrut. Knowledge organization and terminology: philosophical and linguistic bases. **International classification**, v. 19, n. 2, p. 65-71, 1992. DOI: <https://doi.org/10.5771/0943-7444-1992-2-65>.
- DAHLBERG, Ingetrut. Knowledge Organization: its scope and possibilities. **Knowledge Organization**, v. 20, n. 4, p. 211-222, 1993.
- DAHLBERG, Ingetrut. Knowledge organization: a new science?. **Knowledge Organization**, v. 33, n. 1, p. 11-19, 2006.

DAHLBERG, Ingetraut. Concepts & terms: ISKO's major challenge. **Knowledge Organization** v. 36, p. 169-77, 2009.

DAY, Ronald E. **The modern invention of information**: Discourse, history, and power. SIU Press, 2001.

DIAS, Célia da Consolação. A Análise de Domínio, as comunidades discursivas, a garantia de literatura e outras garantias. *Informação & Sociedade: Estudos*, v. 25, n. 2, p. 17, 2015. Disponível em: <<http://hdl.handle.net/20.500.11959/brapci/92174>>. Acesso em: 11 set. 2017.

DODEBEI, Vera Lúcia Doyle. **Tesouro**: linguagem de representação da memória documentária. Niterói: Intertexto; Rio de Janeiro: Interciência, 2002.

ENGERER, Volkmar. Control and syntagmatization: Vocabulary requirements in information retrieval thesauri and natural language lexicons. **Journal of the Association for Information Science and Technology**, v. 68, n. 6, p. 1480-1490, 2017.

ENGINEERS JOINT COUNCIL; Project Lex. **Thesaurus of engineering and scientific terms**. Defense Technical Information Center, 1969.

ESTEBAN-NAVARRO, Miguel Angel; GARCÍA-MARCO, Francisco Javier. Las primeras jornadas sobre organización del conocimiento: organización del conocimiento e información científica. **Scire**: representación y organización del conocimiento, p. 149-157, 1995. Disponível em: <https://ibersid.eu/ojs/index.php/scire/article/view/1038>. Acesso em: 14 abr. 2020.

FEINBERG, Melanie. Hidden bias to responsible bias: an approach to information systems based on Haraway's situated knowledges. **Information Research**, v. 12, n. 4, 2007.

FOGL, Jiri. Relations of the concepts 'information' and 'knowledge'. **International Fórum on Information and Documentation**, The Hague, v.4, n.1, p. 21-24, 1979.

FOSKETT, Antony Charles. **The subject approach to information**. Facet Publishing, 1996.

FOUCAULT, Michel. **A ordem do discurso**. São Paulo: Edições Loyola, 1996.

FREITAS, Juliana Lazzarotto; BUFREM, Leilah Santiago; BREDÁ, Sônia Maria. Methodological choices for research in Information Science: Contributions to domain analysis. **Transinformação**, v. 28, p. 05-14, 2016.

GARCEZ, Dirnele Carneiro; SALES, Rodrigo de. Decolonizando a organização do conhecimento: um olhar do periódico Knowledge Organization (2000-2020). **Tendências da pesquisa brasileira em Ciência da Informação, ANCIB**, v. 14, 2021.

GARCÍA-MARCO, Francisco Javier. The evolution of thesauri and the history of knowledge organization: between the sword of mapping knowledge and the wall of keeping it simple. **Brazilian Journal of Information Science**, v. 10, n. 1, p. 1-11, 2016.

GARCÍA-MARCO, Francisco-Javier. La nueva norma de tesauros UNE ISO 25964-1. **Anuario ThinkEPI**, v. 8, p. 159-163, 2014.

GARDIN, Jean-Claude. Document analysis and linguistic theory. **Journal of Documentation**, v. 29, n. 2, p. 137-168, 1973.

GARDIN, Jean-Claude. Éléments d'un modèle pour la description des lexiques documentaires, **Bulletin des Bibliothèques de France**, n. 05, p. 171-182, 1966. Disponível em: <https://bbf.enssib.fr/consulter/bbf-1966-05-0171-001>. Acesso em: 23 nov. 2016.

GAROZ, Alexei Zaldua. El análisis del discurso en la organización y representación de la información-conocimiento: elementos teóricos. **Acimed**, v. 14, n. 3, p. 1-16, 2006.

- GIL, Antonio Carlos. **Métodos e técnicas de pesquisa social**. 6. ed. Editora Atlas SA, 2008.
- GLÄNZEL, Wolfgang. **Bibliometrics as a research field: a course on theory and application of bibliometric indicators**. Bélgica: [s.n.], 2003. Disponível em https://www.cin.ufpe.br/~ajhol/futuro/references/01%23_Bibliometrics_Module_KUL_BIBLIOMETRICS%20AS%20A%20RESEARCH%20FIELD.pdf. Acesso em: 2 nov. 2021.
- GOMES, Dionei Moreira; FERREIRA, Tânia Borges. Equivalência e definição no Dicionário Bilíngue Português-Mundurukú da área do Magistério: contribuindo para a epistemologia terminológica e terminográfica. **Tradterm**, v. 24, p. 339-364, 2014.
- GOMES, Hagar Espanha. Terminologia e estrutura conceitual. **PontodeAcesso**, v. 15, n. 3, 2021.
- GOMES, Hagar Espanha; MARINHO, Marcílio Teixeira. **Introdução ao estudo do cabeçalho de assunto**. BITI-Biblioteconomia, Informação & Tecnologia da Informação, 1984.
- GOMES, Romeu. Análise e interpretação de dados de pesquisa qualitativa. In: MINAYO, Maria Cecília de Souza; DESLANDES, Suely Ferreira; GOMES, Romeu (Orgs.). **Pesquisa social: teoria, método e criatividade**. 28. ed. Petrópolis: Vozes, 2009, p. 79-108.
- GONZÁLEZ DE GÓMEZ, Maria Nélide. O Domínio das informações em Saúde. In: PINTO, Virginia Bentes; CAMPOS, Henry de Holanda. **Diálogos paradigmáticos: sobre informação para a área da saúde**. Fortaleza: UFC, 2013.
- GONZALEZ, Léia; ORTIZ, María Camila. La categoría político-cultural de la amefricanidad. **Revista Epistemologias do Sul**, v. 6, n. 1, 2022.
- GRÁCIO, Maria Cláudia Cabrini. A coplamente bibliográfico e análise de cocitação: revisão teórico-conceitual. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, v. 21, n. 47, p. 82-99, 2016.
- GRÁCIO, Maria Cláudia Cabrini. **Análises relacionais de citação para a identificação de domínios científicos: uma aplicação no campo dos Estudos Métricos da Informação no Brasil**. Editora UNESP, 2020.
- GRÁCIO, Maria Cláudia Cabrini; OLIVEIRA, Ely Francina Tannuri de. **Análise de cocitação de autores: um estudo teórico-metodológico dos indicadores de proximidade, aplicados ao GT7 da ANCIB**. LIINC em Revista, p. 196-213, 2013.
- GRAF, Ann M. Domain analysis applied to online graffiti art image galleries to reveal knowledge organization structures used within an outsider art community. **Knowledge Organization**, v. 47, n. 7, p. 543-557, 2021.
- GREGOLIN, Maria Do Rosário Valencise. A análise do discurso: conceitos e aplicações. **ALFA: Revista de Linguística**, v. 39, p. 13-21, 1995. Disponível em: <https://periodicos.fclar.unesp.br/alfa/article/view/3967>. Acesso em: 10 abr. 2022.
- GUIMARÃES, José Augusto Chaves et al. A dimensão conceitual da organização do conhecimento no universo científico da ISKO: uma análise de domínio a partir dos congressos de ISKO-Brasil, ISKO-Espanha, ISKO-América Do Norte e ISKO-França. **Scire: representación y organización del conocimiento**, p. 13-26, 2015.
- GUIMARÃES, José Augusto Chaves. Organização do conhecimento: passado, presente e futuro sob a perspectiva da ISKO. **Informação & Informação**, v. 22, n. 2, p. 84-98, 2017. DOI: 10.5433/1981-8920.2017v22n2p84. Acesso em: 3 ago. 2017.

- GUIMARÃES, José Augusto Chaves; TOGNOLI, Natália Bolfarini. Provenance as a domain analysis approach in archival knowledge organization. **Knowledge Organization**, v. 42, n. 8, p. 562-569, 2015.
- GUTIÉRREZ, Antonio Luis García. **Estructura lingüística de la documentación: teoría y método**. EDITUM, 1990.
- HARIRI, Nadjla et al. Evaluation of Persian thesauri of basic sciences based on ANSI/NISO 39.19-2005 standard. **Malaysian Journal of Library & Information Science**, v. 18, n. 1, 2013.
- HARTEL, Jenna. Managing documents at home for serious leisure: a case study of the hobby of gourmet cooking. **Journal of Documentation**, v. 66, n. 6, p. 847-874, 2010.
- HERNANDEZ, Bianca Ferreira; VITAL, Luciane Paula. Estudos de gêneros para a organização do conhecimento no campo da arquivologia. **Informação & Informação**, v. 27, n. 1, p. 1-25, 2022.
- HESSEN, Johannes. Teoria do conhecimento. Martins Fontes. São Paulo, 2012.
- HJØRLAND, B. The classification of Psychology: A case study in the classification of a knowledge field. **Knowledge Organization**, v.25, n.4, p.162-201, 1998.
- HJØRLAND, Birger. Domain analysis in information science: eleven approaches—traditional as well as innovative. **Journal of Documentation**, v. 58, n. 4, p. 422-462, 2002a.
- HJØRLAND, Birger. Epistemology and the socio-cognitive perspective in information science. **Journal of the American Society for Information science and Technology**, v. 53, n. 4, p. 257-270, 2002b.
- HJØRLAND, Birger. Domain Analysis: A Socio-Cognitive Orientation for Information Science Research. **Bulletin of the Association for Information Science and Technology**, v. 30, n. 3, p. 17-21, 2004a.
- HJØRLAND, Birger. Arguments for philosophical realism in library and information science. **Library Trends**, v. 52, n. 3, p. 488–506, 2004b.
- HJØRLAND, Birger. Empiricism, rationalism and positivism in library and information science. **Journal of Documentation**, v. 61, n. 1, p. 130-155, 2005.
- HJØRLAND, Birger. What is knowledge organization (KO)?. **Knowledge Organization**, v. 35, n.2/3, p.86-101, 2008.
- HJØRLAND, Birger. Concept theory. **Journal of the American Society for Information Science and Technology**, v. 60, n. 8, p. 1519-1536, 2009.
- HJØRLAND, Birger. Facet analysis: The logical approach to knowledge organization. **Information Processing & Management**, v. 49, n. 2, p. 545-557, 2013a. DOI: <https://doi.org/10.1016/j.ipm.2012.10.001>.
- HJØRLAND, Birger. Citation analysis: A social and dynamic approach to knowledge organization. **Information Processing & Management**, v. 49, n. 6, p. 1313-1325, 2013b.
- HJØRLAND, Birger. Are relations in thesauri “context-free, definitional, and true in all possible worlds”?. *Journal of the Association for Information Science and Technology*, v. 66, n. 7, p. 1367-1373, 2015.
- HJØRLAND, Birger. Knowledge organization. **Knowledge Organization**, v. 43, n. 6, p. 475-484, 2016. Disponível em: <https://www.nomos-elibrary.de/10.5771/0943-7444-2016-6-475/knowledge-organization-ko-jahrgang-43-2016-heft-6?page=1>. Acesso em: 20 abr. 2020.

HJØRLAND, Birger. Domain analysis. **Knowledge Organization**, v. 44, n. 6, p. 436-464, 2017.

HJØRLAND, Birger. Science, Part II: The Study of Science. **Knowledge Organization**, v. 49, n. 4, p. 273-300, 2022.

HJØRLAND, Birger; ALBRECHTSEN, Hanne. Toward a new horizon in information science: Domain-analysis. **Journal of the Association for Information Science and Technology**, v. 46, n. 6, p. 400-425, 1995.

HJØRLAND, Birger; HARTEL, Jenna. Ontological, epistemological and sociological dimensions of domains. **Knowledge Organization**, v. 30, n. 3-4, p. 239-245, 2003.

HOVY, Eduard. Comparing sets of semantic relations in ontologies. In: **The semantics of relationships**. Springer, Dordrecht, 2002. p. 91-110.

HU, Kai et al. A domain keyword analysis approach extending Term Frequency-Keyword Active Index with Google Word2Vec model. **Scientometrics**, v. 114, p. 1031-1068, 2018.

IBEKWE-SANJUAN, Fidelia. Constructing and maintaining knowledge organization tools: a symbolic approach. **Journal of Documentation**, v. 62, n. 2, p. 229-250, 2006.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Instituto Brasileiro de Geografia e Estatística. **Populações, Censo, 2022**. Disponível em: https://censo2022.ibge.gov.br/panorama/?utm_source=ibge&utm_medium=home&utm_campaign=portal. Acesso em: 20 ago. 2023.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO). **ISO 5964-1985**. Guidelines for the establishment and development of multilingual thesauri, 1985.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO). **ISO 25964 -1**: Thesauri and interoperability with other vocabularies, Part 1: Thesauri for information retrieval. ISO, 2011.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO). **ISO 25964 -2**: Thesauri and interoperability with other vocabularies, Part 2: Interoperability with other vocabularies. ISO, 2013.

JECUPÉ, Kaka Werá. **A terra dos mil povos**: história indígena brasileira contada por um índio. São Paulo: Peirópolis, 1998.

JECUPÉ, Kaka Werá. **Tupã Tenondé**: a criação do Universo, da Terra e do Homem segundo a tradição oral Guarani. Peirópolis, 2020.

JONES, Karen Sparck. Some thesauric history. **Aslib Proceedings**, v. 24, n. 7, p. 400-411, 1972. <https://doi.org/10.1108/eb050353>

KÖNIG, Jason. Re-reading Pollux: encyclopaedic structure and athletic culture in Onomasticon Book 3. **The Classical Quarterly**, v. 66, n. 1, p. 298-315, 2016. DOI: <https://doi.org/10.1017/S0009838816000331>.

KRIEGER, Maria da Graça. Das relações entre Lexicologia e Terminologia. In: CATALÀ, Sara Álvarez; BARITÈ, Mario (Orgs.). **Teoría y praxis em Terminología**. Universidad de la República: Montevideo, 2017, p. 151-160.

KRIEGER, Maria da Graça. Terminologia revisitada. **D.E.L.T.A.**, v. 16, n.2, p. 209-228, 2000.

KRIEGER, Maria da Graça; FINATTO, Maria José Bocorny. **Introdução à terminologia**: teoria e prática. São Paulo: Contexto, 2004.

KROOKS, David A.; LANCASTER, F. Wilfrid. The evolution of guidelines for thesaurus construction. **Libri**, v. 43, n.4, p. 326-342, 1993.

KUHN, Thomas S. **A estrutura das revoluções científicas**. 9. ed. São Paulo: Perspectivas, 2005.

LANCASTER, Frederic Wilfrid. **Vocabulary control for information retrieval**. 2.ed. Arlington: Information Resources Press, 1986.

LARA, Marilda Lopes Ginez de. O processo de construção da informação documentária e o processo de conhecimento. **Perspect. cienc. inf.**, Belo Horizonte, v. 7, n. 2, p. 127-139, jul./dez. 2002.

LARA, Marilda Lopes Ginez de. **Representação e linguagens documentárias: bases teórico-metodológicas**. 1999, 208 f. Tese (Doutorado em Ciências da Comunicação) – Departamento de Biblioteconomia e Documentação, Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 1999. Disponível em: <https://www.teses.usp.br/teses/disponiveis/27/27143/tde-02122019-153131/publico/MarildaLopesGinesdeLaraDoutorado.pdf>. Acesso em: 10 fev. 2018.

LEE, Jae Yun; KIM, Heejung; KIM, Pan Jun. Domain analysis with text mining: Analysis of digital library research trends using profiling methods. **Journal of Information Science**, v. 36, n. 2, p. 144-161, 2010.

LÓPEZ, Miriam Tomás. Diseño de un tesauro conceptual de folklore de la Región de Murcia. **Scire: representación y organización del conocimiento**, p. 65-76, 2004.

LÓPEZ-HUERTAS, María J. Domain analysis for interdisciplinary knowledge domains. **Knowledge Organization**, v. 42, n. 8, p. 570-580, 2015. DOI: doi.org/10.5771/0943-7444-2015-8-570.

LYONS, John. **Lingua(gem) e linguística: uma introdução**. Rio de Janeiro: LTC, 2011.

MAI, Jens-Erik. Analysis in indexing: document and domain centered approaches. **Information processing & management**, v. 41, n. 3, p. 599-611, 2005.

MAI, Jens-Erik. Actors, domains, and constraints in the design and construction of controlled vocabularies. **Knowledge Organization**, v. 35, n. 1, p. 16-29, 2008.

MARCONDES, Carlos Henrique. Fundamentos da organização do conhecimento. **PontodeAcesso**, v. 15, n. 3, p. 249-282, dez., 2021.

MARTINES, Alexandre Robson; ALMEIDA, Carlos Cândido. A linguagem documentária e o signo: uma análise sobre termo e conceito. **Informação & Informação**, v. 26, n. 3, p. 697-728, 2021.

MARTINES, Alexandre Robson; MOREIRA, Walter; DE ALMEIDA, Carlos Cândido. Do signo ao tesauro: contribuições de três correntes da linguagem. **Ciência da Informação**, v. 51, n. 1, 2022.

MARTÍNEZ DE SOUSA, José. **Diccionario de biblología y ciencias afines**. Madrid: Fundación Germán Sánchez Ruipérez, 1989.

MARTÍNEZ-ÁVILA, Daniel; GUIMARÃES, José Augusto Chaves; EVANGELISTA, Isadora Victorino. Epistemic communities in Knowledge Organization: An analysis of the NASKO meetings proceedings. **NASKO**, v. 6, n. 1, p. 113-123, 2017.

MARTÍNEZ-GONZÁLEZ, M. Mercedes; ALVITE-DÍEZ, María-Luisa. A semantic web methodological framework to evaluate the support of integrity in thesaurus tools. *Journal of Information Science*, v. 46, n. 3, p. 378-391, 2020.

MARTINS, Gracy Kelli. CÔRTEZ, Gisele Rocha. A representação da Informação e do Conhecimento nas Representações Sociais: intersecções e limites. In: ALBUQUERQUE, Maria Elizabeth Baltar Caneiro; MARTINS, Gracy Kelli; MOTA, Denyson Axel Ribeiro (orgs.) **Organização e representação da informação e do conhecimento: intersecções teórico-sociais**. João Pessoa: Editora UFPB, 2019, p. 159-182.

MARTINS, Gracy Kelli; MORAES, João Batista Ernesto de. Análise de Domínio: uma proposta metodológica para extração de indicadores cognitivo e social de institucionalização científica. In: PINTO, Virgínia Bentes; VIDOTTI, Silvana Aparecida Borsetti Gregório; CAVALCANTE; Lúcia Eugênia. **Aplicabilidades metodológicas em Ciência da Informação**. Fortaleza, Edições UFC, 2015, p. 73-100.

MAZZOCCHI, Fulvio et al. **Relational semantics in thesauri**: some remarks at theoretical and practical levels. *Knowledge Organization*, v. 34, n. 4, p. 197-214, 2007.

MELO, Fabio J. Dantas; BRÄSCHER, Marisa. **Fundamentos da linguística para a formação do profissional de informação**. Brasília: Centro Editorial, 2011.

MÉNARD, Elaine; MAS, Sabine; ALBERTS, Inge. Faceted classification for museum artefacts: A methodology to support web site development of large cultural organizations. In: Aslib Proceedings. **Emerald Group Publishing Limited**, 2010. p. 523-532.

MENDONÇA, Renato. **A influência africana no português do Brasil**. Brasília : FUNAG, 2012.

MEY, Eliane Serrão Alves; SILVEIRA, Naira Chistofelett. Breve histórico dos catálogos e da catalogação. In: **Catalogação no plural**. Brasília: Brinquet de Lemos, 2009, p. 59-89.

MIGNOLO, Walter D. A Geopolítica do Conhecimento e a Diferença Colonial. **Revista Lusofona de Educação**, v. 48, n. 48, p. 187-224, 2020. Disponível em: <https://doi.org/10.24140/issn.1645-7250.rle48.12>. Acesso em: 20 fev. 2023.

MILES, Alistair; BECHHOFFER, Sean. **SKOS simple knowledge organization system reference**. W3C recommendation, 2009. Disponível em: <https://www.w3.org/TR/skos-reference/>. Acesso em: 24 mar. 2022.

MINAYO, Maria Cecília de Souza. O desafio da pesquisa social. In: MINAYO, Maria Cecília de Souza; DESLANDES, Suely Ferreira; GOMES, Romeu (Orgs.). **Pesquisa social: teoria, método e criatividade**. 28. ed. Petrópolis: Vozes, 2009, p. 9-29.

MORAES, Miriam Gontijo de. Tensão identitária e organização do conhecimento: olhar epistemográfico. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 11, n. 2, 2018.

MORANDO, María Agustina. Los mecanismos de innovación léxica y la reconfiguración de la modernidad indígena en guaraní chaqueño. **Revista del Museo de Antropología**, v. 14, n. 1, p. 173-184, 2021.

MOREIRA, Walter. Tesauros e ontologias como modelos de sistemas de organização do conhecimento. **Brazilian Journal of Information Science: research trends**, v. 13, n. 1, p. 15-20, 2019.

- MOREIRO, José Antonio. Adaptación de los vocabularios documentales al ambiente digital en red: léxico, significado y relaciones semánticas. **Informação & Sociedade**, v. 28, n. 1, p. 35-46, jan./abr., 2018.
- MOTTA, Dilza Fonseca. Modelo relacional como nova abordagem para a construção de tesouros. Rio de Janeiro: **Biblioteconomia, Informação & Tecnologia da Informação**, 1987. Disponível em <http://www.conexaorio.com/bit/>. Acesso em: 1 nov. 2021.
- MOTTA, Marta; TIBERI, Melissa. Riflessi dello standard britannico BS 8723 nel Thesaurus del Nuovo Soggettario. **Bollettino AIB (1992-2011)**, v. 49, n. 3, p. 325-342, 2009.
- MOYA-ANEGÓN, Félix; HERRERO-SOLANA, Víctor; JIMÉNEZ-CONTRERAS, Evaristo. A connectionist and multivariate approach to science maps: the SOM, clustering and MDS applied to library and information science research. **Journal of information science**, v. 32, n. 1, p. 63-77, 2006.
- MUNANGA, Kabengele. **Origens africanas do Brasil contemporâneo: histórias, línguas, culturas e civilizações**. São Paulo: Global, 2009.
- MUNANGA, Kabengele. Uma abordagem conceitual das noções de raça, racismo, identidade e etnia. In: BRANDÃO, André Augusto. **Programa de educação sobre o negro na sociedade brasileira**. Niterói: EdUFF, 2004, p. 17-34.
- NASCIMENTO, Denise Morado; MARTELETO, Regina Maria. Social field, domains of knowledge and informational practice. **Journal of Documentation**, v. 64, n. 3, p. 397-412, 2008. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/00220410810867605/full/html>. Acesso em: 10 ago. 2023.
- NATIONAL ARCHIVES, The. **Records of the Royal Radar Establishment, later the Royal Signals and Radar Establishment, and its predecessors** [on-line]: 2023. Disponível em: <https://discovery.nationalarchives.gov.uk/details/r/C550>. Acesso em: 30 abr. 2022.
- NEIGHBORS, James Milne. **Software construction using components**. Tese de Doutorado. Department of Information and Computer Science, University of California, Irvine, 1980.
- NIELSEN, Lykke Marianne. Domain analysis, an important part of thesaurus construction. **Advances in classification research online**, v. 11, n. 1, p. 9-50, 2000.
- NIELSEN, Lykke Marianne. Thesaurus Construction: Key Issues and Selected Readings, *Cataloging & Classification Quarterly*, 37:3-4, 57-74, 2004.
- OLIVEIRA, Dalgiza Andrade; ARAUJO, Ronaldo Ferreira de. 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**, v. 17, n. 34, p. 17-30, 2012.
- OLIVEIRA, Ely Francina Tannuri de; GRÁCIO, Maria Cláudia Cabrini. Studies of author cocitation analysis: a bibliometric approach for domain analysis. **Iris**, Recife, v. 2, n. 1, p. 12-23, 2013.
- OLSON, Hope A. A potência do não percebido: Hegel, Dewey e seu lugar na corrente principal do pensamento classificatório. **InCID**, v. 2, n. 1, p. 3-15, 2011.
- OLSON, Hope A. How we construct subjects: A feminist analysis. **Library trends**, v. 56, n. 2, p. 509-541, 2007.
- OLSON, Hope A. **The power to name: locating the limits or subject representation in libraries**. Dordrecht: Kluwer Academic Publisher, 2002.

OWENS, Leslie Ann; COCHRANE, Pauline Atherton. Thesaurus evaluation. **Cataloging & classification quarterly**, v. 37, n. 3-4, p. 87-102, 2004.

PADOVANI, Bruna Fernanda Soares de Lima. **Estudo do léxico da língua Apurinã uma proposta de macro e microestrutura para o dicionário Apurinã**. 2020. 2 v. Tese (Doutorado em Letras) - Instituto de Letras e Comunicação, Universidade Federal do Pará, Belém, 2020. Disponível em: <http://repositorio.ufpa.br:8080/jspui/handle/2011/15372>. Acesso em: 20 out. 2023.

PAIS, Cidmar Teodoro. Aspectos de uma tipologia dos universos de discurso. São Paulo, **Revista Brasileira de Linguística**, v. 6, n. 1, p. 43-65, 1984.

PAJÍĆ, Vesna et al. Semi-automatic extraction of multiword terms from domain-specific corpora. **The Electronic Library**, v. 36, n. 3, p. 550-567, 2018.

PASTOR-SÁNCHEZ, Juan-Antonio. Proposal to represent the UNESCO Thesaurus for the semantic web applying ISO-25964. **Brazilian Journal of Information Studies: Research Trends**, v. 10, n. 1, p. 1-8, 2016.

PEÑA, Gilberto Anguiano; PEÑA, Catalina Naumis. Extraction of candidate terms from a corpus of non-specialized, general language. **Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información**, v. 29, n. 67, p. 19-45, 2015.

PETTER, Margarida Maria Taddoni. **Línguas africanas no Brasil**. Gragoatá, v. 10, n. 19, 2005.

PINHO, Fábio Assis; GUIMARÃES, José Augusto Chaves. Os desafios da representação do conhecimento face à homossexualidade masculina. In: **Desafios e perspectivas científicas para a organização e representação do conhecimento na atualidade ISKO Brasil**, v. 1, p. 143-146, 2011. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/135003>. Acesso em: 14 jul. 2022.

PRANDI, Reginaldo. De africano a afro-brasileiro: etnia, identidade, religião. **Revista USP**, n. 46, p. 52-65, 2000.

PRIETO-DIAZ, Rubin. Domain Analysis: An Introduction, ACM SIGSOFT Software. **Engineering Notes**, v. 15, p.47-54, abr., 1990.

QUESADA, Manuel Alejandro Romero; QUINTANA, Ania Rosa Hernández. El método etnográfico y su relación con el análisis de dominio. **Biblios Journal of Librarianship and Information Science**, n. 61, p. 70-84, 2015.

QUIJANO, Aníbal. Colonialidad del poder, cultura y conocimiento en América Latina. **Dispositio**, v. 24, n. 51, p. 137-148, 1999.

QUIJANO, Aníbal. “Raza”, “etnia” y “nación” en Mariátegui: cuestiones abiertas. In: QUIJANO, Aníbal. **Cuestiones y horizontes: de la dependencia histórico-estructural a la colonialidad/descolonialidad del poder**. Buenos Aires: CLACSO, 2014, p. 757-775.

RAFFESTIN, Claude. Raça, etnia e poder. In: RAFFESTIN, Claude. **Por uma geografia do poder**. São Paulo: Ática, 1993, p. 130-139.

RAMALHO, Rogério Aparecido Sá. Análise do modelo de dados SKOS: Sistema de Organização do Conhecimento Simples para a Web. **Informação & Tecnologia**, v. 2, n. 1, p. 66-79, 2015. Disponível em: <https://periodicos.ufpb.br/ojs/index.php/itec/article/view/25995>. Acesso em: 11 abr. 2020.

RANGANATHAN, Shiyali Ramamrita. **Prolegomena to library classification**. 3 ed. Bombaim: Asia Publishing House, 1967.

RANGANATHAN, Shiyali Ramamrita. **Prolegomena to library classification**. Madras: Madras Library Association, 1937. (Séries 6)

RAUCH, Mirta; BARITÉ, Mario. Glosário de Terminología. In: CATALÁ, Sara Álvarez; BARITÉ, Mario (Org.). **Teoría y praxis em terminologia**. Uruguay: CSIC; Universidad de la Republica Uruguay, 2017.

REES-POTTER, Lorna K. Dynamic thesaural systems: A bibliometric study of terminological and conceptual change in sociology and economics with application to the design of dynamic thesaural systems. **Information Processing & Management**, v. 25, n. 6, p. 677-689, 1989.

REES-POTTER, Lorna K. Dynamic thesaural systems: A bibliometric study of terminological and conceptual change in sociology and economics with application to the design of dynamic thesaural systems. **Information Processing & Management**, v. 25, n. 6, p. 677-689, 1989.

REIS, Daniela et al. Tratamento descritivo e temático da informação: recomendações para estudos sobre aspectos semióticos na criação de registros bibliográficos. **Revista Ibero-Americana de Ciência da Informação**, v. 11, p. 42-58, 2018.

REIS, Mateus da Silva; BARROS, Thiago Henrique Bragato; SANTOS JUNIOR, Roberto Lopes dos. Proposta metodológica para a elaboração do microtesauro do acervo de medalhística do Museu do Instituto Histórico e Geográfico do Pará. **Memória e Informação**, v. 5, n. 1, p. 65-75, 2021.

RIBEIRO, Djamila. **Lugar de fala**. São Paulo: Pólen, 2019.

RIGO, Sandro José; DA ROSA ALVES, Isa Mara; BARBOSA, Jorge Luis Victória. A linguistic approach to identify the affective dimension expressed in textual messages. **International Journal of Information and Communication Technology Education (IJICTE)**, v. 11, n. 1, p. 44-59, 2015.

ROBERTS, Norman. The pre-history of the information retrieval thesaurus. **Journal of Documentation**, v. 40, n.4, p. 271-285, 1984. <https://doi.org/10.1108/eb026769>.

RODRIGUES, Anderson Luiz Cardoso. A complexidade da cultura amazônica e seu reflexo para a organização e representação da informação. **AtoZ: novas práticas em informação e conhecimento**, v. 1, n. 2, p. 10-25, 2012.

RODRIGUES, Anderson Luiz Cardoso. A cultura e a organização do conhecimento: desafios teórico-metodológicos. **Información, cultura y sociedad**, n. 32, p. 37-58, 2015.

RODRIGUES, Aryon Dall'Igna. **Línguas brasileiras**: para o conhecimento das línguas indígenas. São Paulo: Loyola, 1986.

RODRIGUES, Aryon Dall'Igna. **Línguas indígenas**: 500 anos de descobertas e perdas. DELTA: Documentação E Estudos Em Linguística Teórica E Aplicada, v. 9, n. 1, 1993.

RODRIGUES, Aryon Dall'Igna. Sobre as línguas indígenas e sua pesquisa no Brasil. **Ciência e cultura**, v. 57, n. 2, p. 35-38, 2005.

RODRIGUES, Raymundo Nina. **Os últimos africanos**: nações pretas que se extinguem. In: Os africanos no Brasil. 2 ed. São Paulo: Companhia Editora Nacional, 1935, p. 151-184.

ROSENFELD, Louis; MORVILLE, Peter; ARANGO, Jorge. **Information architecture**: for the web and beyond. O'Reilly, 2015.

- ROUGÉ, Jean-Louis. A inexistência de crioulo no Brasil. In: PETTER, Margarida; FIORIN, José Luiz. **África no Brasil: a formação da língua portuguesa**. São Paulo: Contexto, 2014, p. 63-74.
- SALDANHA, Gustavo Silva; DE SOUZA, Rosali Fernandez. Teoria barroca da organização do conhecimento: Emanuele Tesauro e o espelho turvo das tensões entre epistemologia, metodologia e sociedade. **Informação & Informação**, v. 22, n. 2, p. 11-32, 2017.
- SAMPIERI, Roberto Hernández; COLLADO, Carlos Fernández; LUCIO, María del Pilar Baptista. **Metodologia de pesquisa**. Porto Alegre: Penso, 2013.
- SANCHEZ-CUADRADO, Sonia; COLMENERO-RUIZ, Maria-Jesus; MOREIRO, Jose-Antonio. Thesauri: standards and guidelines. **El profesional de la información**, v. 21, n. 3, p. 229-235, 2012.
- SARVABHOTLA, Kiran; PINGALI, Prasad; VARMA, Vasudeva. Sentiment classification: a lexical similarity based approach for extracting subjectivity in documents. **Information Retrieval**, v. 14, p. 337-353, 2011.
- SCHMITZ, John Robert. O léxico: campo minado, área conflituosa, palco de sedução e depósito de recortes. In: ALVES, Ieda Maria; JESUS, Ana Maria (Orgs.). **Os estudos lexicais em diferentes perspectivas** [livro eletrônico]: volume V. São Paulo : FFLCH/USP, 2015, p. 9-28.
- SCHNEIDER, Jesper W. Verification of bibliometric methods' applicability for thesaurus construction. In: **ACM SIGIR Forum**. New York, NY, USA: ACM, 2005. p. 63-64.
- SCHNEIDER, Jesper W.; BORLUND, Pia. Introduction to bibliometrics for construction and maintenance of thesauri: Methodical considerations. **Journal of Documentation**, v. 60, n. 5, p. 524-549, 2004.
- SCHNEIDER, Jesper W.; BORLUND, Pia. Preliminary study of the potentiality of bibliometric methods for the construction of thesauri. In: **Emerging Frameworks and Methods: Proceedings of the Fourth International Conference on Conceptions of Library and Information Science (CoLIS 4)**, Seattle, WA. 2002. p. 151-165.
- SCHNEIDER, Jesper. Concept symbols revisited: Naming clusters by parsing and filtering of noun phrases from citation contexts of concept symbols. **Scientometrics**, v. 68, n. 3, p. 573-593, 2006.
- SHIRI, Ali Asghar; REVIE, Crawford; CHOWDHURY, Gobinda. Thesaurus-assisted search term selection and query expansion: a review of user-centred studies. **Knowledge Organization**, v. 29, n. 2, p. 1-19, 2002.
- SHUTER, Janet. New and revised indexing standards. **The Indexer: The International Journal of Indexing**, v. 26, n. 2, p. 96-98, 2008.
- SILVA, Franciéle Carneiro Garcês da; SALDANHA, Gustavo Silva. Biblioteconomia negra brasileira: caminhos, lutas e transformação. **Tendências da pesquisa brasileira em Ciência da Informação**, v. 12, n. 2, p. 1-24, 2019.
- SILVA, Franciéle Carneiro Garcês da; ROMEIRO, Nathália Lima. **O protagonismo da mulher na Arquivologia, Biblioteconomia, Museologia e Ciência da Informação**. Florianópolis: Rocha Gráfica e Editora, 2019.
- SILVA, Franciéle Carneiro Garcês. Perspectivas críticas e epistemologias negras na Biblioteconomia. In: CARDONA, Natalia Duque; SILVA, Franciéle Carneiro Garcês. **Epistemologias latino-americanas na Biblioteconomia e Ciência da informação**:

Contribuições da Colômbia e do Brasil. Florianópolis: Rocha Gráfica e Editora, 2020, p.73-118.

SILVA, Jonathas Luiz Carvalho; FARIAS, Maria Giovanna Guedes. Reflexões teóricas sobre a construção paradigmática da Ciência da Informação: considerações acerca do(s) paradigma(s) cognitivo(s) e social. **Biblios**: Revista de Bibliotecología y Ciencias de la Información, Pittsburgh, n. 51, p. 42-56, 2013.

SMIRAGLIA, Richard P. Domain coherence within knowledge organization: people, interacting theoretically, across geopolitical and cultural boundaries. In: **Proceedings of the Annual Conference of CAIS/Actes du congrès annuel de l'ACSI**, 2011.

SMIRAGLIA, Richard P. Epistemology of domain analysis. In: Richard P. Smiraglia and Hur-Li Lee. (Ogr.). **Cultural frames of knowledge**. Würzburg: Ergon Verlag, p. 111-124, 2012.

SMIRAGLIA, Richard P. **The elements of knowledge organization**. Cham, Springer, 2014.

SMIRAGLIA, Richard P. The epistemological dimension of knowledge organization. **Iris**, v. 2, n. 1, p. 2-11, 2013.

SMIRAGLIA, Richard P. **Domain Analysis for Knowledge Organization**: Tools for Ontology Extraction. Chandos Information Professional Series. Oxford: Elsevier/Chandos, 2015a.

SMIRAGLIA, Richard P. Domain Analysis of Domain Analysis for Knowledge Organization: Observations on an Emergent Methodological Cluster. **Knowledge Organization**, v. 42, n. 8, p. 602-611, 2015b.

SMIRAGLIA, Richard P.; LÓPEZ-HUERTAS, María J. Domain analysis redux: An introduction. **Knowledge Organization**, v. 42, n. 8, p. 553-556, 2015.

SOERGEL, Dagobert. Thesauri and ontologies in digital libraries. In: **International Conference on Digital Libraries**: Proceedings of the 5 th ACM/IEEE-CS joint conference on Digital libraries, 2005.

SOMERS, H. L. Observations on standards and guidelines concerning thesaurus construction. **Knowledge Organization**, v. 8, n. 2, p. 69-74, 1981.

SONG, Yanhui et al. Studying domain structure: a comparative analysis of bibliographic coupling analysis and co-citation analysis considering all authors. **Online Information Review**, v. 47, n. 1, p. 123-137, 2023.

SOUZA, Jacqueline A.; PESTANA, Olivia. Elos interdisciplinares para estruturação semântica com vistas à organização da informação. **Páginas a&b**: arquivos e bibliotecas, p. 245-257, 2017.

STOCK, Wolfgang G. Concepts and semantic relations in information science. **Journal of the American Society for Information Science and Technology**, v. 61, n. 10, p. 1951-1969, 2010.

SVENONIUS, Elaine. Subject catalogues: Headings and structure: EJ Coates. Library Association, London (1988). **Informalion Processing & Managemenf**, v. 26, n. 1, p. 187-195. 1990.

SWALES, John M. The concept of discourse community. In: SWALES, John M. **Genre analysis**. Cambridge: Cambridge University Press, 1990, p.21-32.

TENNIS, Joseph T. Com o que uma Análise de Domínio se parece no tocante a sua forma, função e gênero?. **Brazilian Journal of Information Science**, Marília, v. 6, n. 1, p. 3-15, 2012.

TENNIS, Joseph T. Two axes of domains for domain analysis. **Knowledge Organization**, v. 30, n. 3/4, p. 191-195, 2003.

TESSER, Gelson João. Principais linhas epistemológicas contemporâneas. **Educar em revista**, n. 10, p. 91-98, 1995.

THELLEFSEN, Martin Muderspach. Domain analytical information and knowledge organization: investigating the externalist and internalist conception of information. **Journal of Documentation**, v. 79, n. 1, p. 21-35, 2023.

TOMASZUK, Dominik. CoVoMe: New Methodology for Building Controlled Vocabulary. In: **Research Conference on Metadata and Semantics Research**. Cham: Springer International Publishing, 2021. p. 41-56.

URDICIÁIN, Blanca Gil. **Evolución histórica de los tesauros españoles y análisis de su rendimiento en el proceso de recuperación de información**. 1997. 429 f. Tese de Doutorado – Facultad de Ciencias de la Información, Departamento de Biblioteconomía y Documentación, Universidad Complutense de Madrid, Madrid, 1997.

VAN DIJK, Teun A. **Discurso e poder**. 2 ed. São Paulo: Contexto, 2010.

VICKERY, Brian C. Thesaurus: a new word in documentation. **Journal of Documentation**, v. 16, n. 4, dec. 1960.

VICKERY, Brian. Faceted classification for the web. **Axiomathes**, v. 18, n. 2, p. 145-160, 2008. DOI: <https://doi.org/10.1007/s10516-007-9025-9>.

WANG, Lin; QIU, Junping. Domain analytic paradigm: a quarter century exploration of fundamental ideas in information science. **Journal of Documentation**, v. 78, n. 5, p. 1027-1052, 2021. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/JD-12-2020-0219/full/html>. Acesso em: 10 jul. 2023.

WANG, Shengang. The intellectual landscape of the domain of culture and ethics in knowledge organization: an analysis of influential authors and works. **Cataloging & Classification Quarterly**, v. 57, n. 4, p. 227-243, 2019.

WEISS, Leila Cristina; BRÄSCHER, Marisa. Relações semânticas em tesauros: contribuições da abordagem pragmática. **InCID: Revista de Ciência da Informação e Documentação**, v. 7, n. 2, p. 136-155, 2016.

WHITLEY, Richard. Cognitive and social institutionalization of scientific specialities and research areas. In: WHITLEY, Richard. **Social processes of scientific development**. London: Routledge and Kegan, 1974. p. 69-95.

WILL, Leonard. The ISO 25964 data model for the structure of an information retrieval thesaurus. **Bulletin of the American Society for Information Science and Technology**, v. 38, n. 4, p. 48-51, 2012.

WRIGHT, Kieth C. Trends in Modern Subject Analysis with Reference to Text Derivative Indexing and Abstracting Methods: The State of the Art. Information-Part 2. **ERIC Clearing house on Library and Information Sciences**, Washington, set./out., 1972. Disponível em: <https://eric.ed.gov/?id=ED073787>. Acesso em: 20 dez. 2021.

WÜSTER, Eugen. **Introducción a la teoría general de la terminología ya la lexicografía terminológica**. Barcelona: IULA, 1998.

YOON, JungWon. Towards a user-oriented thesaurus for non-domain-specific image collections. **Information Processing & Management**, v. 45, n. 4, p. 452-468, 2009.

ZAMUDIO, Rebeca Mariana Gaytán. Violencia epistémica y creación de subjetividades coloniales. In: LÓPEZ NÁJERA, Verónica Renata. (Coord.). **De lo poscolonial a la descolonización: Genealogías latinoamericanas**. Ciudad de México: Universidad Nacional Autónoma de México, 2018. 218 p.

ZENG, Marcia Lei. Knowledge organization systems (KOS). **Knowledge Organization**, v. 35, n. 2-3, p. 160-182, 2008. DOI: 10.5771/0943-7444-2008-2-3-160. Acesso em: 20 ago. 2017.

ZHOU, Ning; CHEN, Xuyi; WANG, Jianhai. Study on Methods of Knowledge Domain Visualization. In: **Proceedings of International Conference On Innovation Management, 2009, Wuhan**. Anais [...]. Los Alamitos: IEEE Computer Society, 2009. p.15-18.

APÊNDICE A - FICHA DE REGISTRO DOCUMENTAL

Título do artigo: Autores: Ano:
Resumo Objetivos: Metodologia/Métodos: Resultados: Conclusão:
Eixo I – Aspectos conceituais 1.1 Conceitos 1.2 Definições dos conceitos 1.3 Pontos conflitantes 1.4 Proposta
Eixo II – Aspectos teóricos 2.1 Teoria 2.2 Paradigma 2.3 Abordagem teórica 2.4 Epistemologias 2.5 Pontos conflitantes 2.6 Proposta
Eixo III – Aspectos metodológicos 3.1 Abordagem metodológica 3.2 Métodos 3.3 Ferramentas/instrumentos 3.4 Padrões 3.5 Diretrizes 3.6 Pontos conflitantes 3.7 Proposta

APÊNDICE B – DOCUMENTOS DA PESQUISA EMPÍRICA

Autores	Ano	Título
Alencar, Maira Fernandes; Cervantes, Brigida Maria Nogueira	2021	Looking at the epistemological dimension of Brazilian knowledge organization with three lenses of domain analysis: epistemological, scientific structuring and terminology
Albrechtsen, Hanne	2015	This is Not Domain Analysis.
Albuquerque, Ana Cristina	2023	Por um reflexão sobre a organização e representação de conceitos decoloniais na américa latina: O pensamento de Aníbal Quijano à luz da Análise de Domínio
Alencar, Maira Fernandes; Nogueira Cervantes, Brigida Maria	2019a	Access to archival information: a theoretical approach of the post-custodial strand with the functional thesaurus
Alencar, Maira Fernandes; Nogueira Cervantes, Brigida Maria	2019b	Organization and representation of the archival knowledge: in pursuit of a method to construct a functional thesaurus
Alexiev, Vladimir; Isaac, Antoine; Lindenthal, Jutta	2016	On the composition of ISO 25964 hierarchical relations (BTG, BTP, BTI).
Almeida, Josiana Florêncio Vieira Régis de; Dias, Guilherme Ataíde	2019	Estado da arte sobre análise de domínio no campo da ciência da informação brasileira
Almeida, Maurício Barcellos; Emygdio, Jeanne Louize	2019	Uma investigação teórica sobre relações semânticas partitivas e sua aplicação em sistemas de organização do conhecimento
Alvarenga, Lídia; Dias, Célia da Consolação	2012	Análise de Domínio e Gestão Arquivística
Amorim, Igor Soares; Café, Lígia Maria	2017	Agenciamento e Análise de Domínio: um encontro possível
Amorim, Igor Soares; Vianna, William Barbosa; Brasher, Marisa	2019	Aspectos epistemológicos em estudos sobre análise de domínio sob as lentes do racionalismo e do empirismo
Araújo, Dayanne Albuquerque; Cavati Sobrinho, Heliomar	2020	Tabli: protótipo de aplicativo gamificado para a construção de tesouros
Arencibia-Jorge, Ricardo; Vega-Almeida, Rosa Lidia; Martí-Lahera, Yohannis	2007	Domain Analysis for the Construction of a Conceptual Structure: A Case Study
Barite, Mario; Fernandez-Molina, Juan Carlos; Chaves Guimaraes, Jose Augusto; Ernesto de Moraes, Joao Batista	2010	Literary warrant: elements for a critical review after one century
Barros, Thiago Henrique Bragato	2023	Michel Pêcheux's Discourse Analysis: An Approach to Domain Analyses.
Beghtol, Clare	1995	Domain analysis, literary warrant, and consensus - the case of fiction studies
Bentes Pinto, Virginia; Campos, Henry de Holanda; Aragão, Maria do Socorro; Ribeiro, Cristina da Costa	2020	O léxico das comunidades indígenas do Ceará na designação de doenças: reflexões para a construção de vocabulário controlado
Bezares, Gonzalo Mochon; Rodriguez, Eva Maria Mendez; La Fuente, Gema Bueno de	2020	Normalized conversion (SKOS) of interoperable knowledge organization systems on the Web

Bissoli, Bruna da Silva; Covello, Lucas Gatto; Pisseli, Bianca Iris; Santos, Raphael Augusto	2018	Identidade de gênero e diversidade sexual: proposta de elaboração de microtesauro
Blanco Correa, Óscar Elías; Salcedo Lagos, Pedro; Kotz Grabole, Gabriela Emilce	2020	Análisis del léxico de las emociones: una aproximación desde la disponibilidad léxica y la Teoría de los Grafos Léxicos
Bonaccorsi, Andrea; Chiarello, Filippo; Fantoni, Gualtiero	2021	Impact for whom? Mapping the users of public research with lexicon-based text mining
Bräscher, Marisa	2014	Semantic Relations in Knowledge Organization Systems
Broughton, Vanda	2006b	The need for a faceted methods of information retrieval
Café, Luísa Chaves; Cavalcanti, Vanessa Oliveira de Macêdo; Barros, Camila Monteiro de; Vital, Luciane Paula; Alvarez, Edgar Bisset	2022	Domain analysis of Brazilian open access institutional policies
Calvitti, Tiziana; Viti, Elisabetta	2009	Da ISO 2788 ai nuovi standard per la costruzione e l'interoperabilità dei vocabolari controllati: un'analisi comparativa.
Caminotti, María Laura; Martínez, Ana María	2006	Fútbol, tesauros y taxonomías Web: desafíos del control de vocabulario.
Campos, Luiz Fernando de Barros; Venâncio, Ludmila Salomão	2007	Perspectivas em (in)formação: tendências e tensões entre abordagens físicas, cognitivistas e emergentes
Campos, Maria Luiza Almeida; Gomes, Hagar Espanha	2005	Princípios de Organização e Representação do Conhecimento na Construção de Hiperdocumentos
Campos, Maria Luiza Almeida; Gomes, Hagar Espanha	2006	Metodologia de Elaboração de Tesauro Conceitual: a categorização como princípio norteador
Carlan, Eliana; Bräscher, Marisa	2011	Sistemas de Organização do Conhecimento na visão da Ciência da Informação
Chen, Sophy Shu-Jiun	2021	Methodological Considerations for Developing the Art & Architecture Thesaurus in the Chinese Language and its Applications
Clarke, Stella G. Dextre	2008	ISO 2788 + ISO 5964 + Much Energy = ISO 25964
Clarke, Stella G. Dextre; Zeng, Marcia Lei	2012	From ISO 2788 to ISO 25964: The Evolution of Thesaurus Standards towards Interoperability and Data Modeling
Coladangelo, L. P.	2020	Thesaurus and Ontology Construction for Contra Dance: Knowledge Organization of a North American Folk Dance Domain
Coladangelo, L. P.	2021	Organizing Controversy: Toward Cultural Hospitality in Controlled Vocabularies Through Semantic Annotation.
Colín, Marisela	2023	Los nombres del racismo en México: una mirada léxica
Damus, Maria Arminda; Acuna, Gabriela Noemi	2019	Approach to Domain-analysis from Library and Information Science research
Feinberg, Melanie	2007	Hidden bias to responsible bias: an approach to information systems based on Haraway's situated knowledges.
Ferreira, Ana Carolina; Maculan, Benildes Coura Moreira dos Santos	2020	Fundamentos teórico-conceituais aplicáveis à revisão e à atualização de tesauros
Freitas, Juliana Lazzarotto; Bufrem, Leilah Santiago; Breda, Sonia Maria	2016	Methodological choices for research in Information Science: Contributions to domain analysis

Fujita, Mariangela Spotti Lopes; Branco, Luciana Beatriz Piovezan Rio; Alves, Roberta Caroline Vesu	2020	Evolução teórica sobre tesauro: estudo diacrônico das publicações de Dagobert Soergel (1972-2017)
Garoz, Alexei Zaldúa	2006	El análisis del discurso en la organización y representación de la información y el conocimiento
Golub, Koraljka; Bergenmar, Jenny; Humelsjo, Siska	2022	Searching for Swedish LGBTQI fiction: challenges and solutions
Gomes, Hagar Espanha	2021	Terminologia e estrutura conceitual
Gomes, Hagar Espanha; Campos, Maria Luiza Almeida	2004	Tesauro e normalização terminológica: o termo como base para intercâmbio de informações
Gomes, Pablo; Frota, Maria Guiomar da Cunha	2019	Knowledge Organization from a Social Perspective: Thesauri and the Commitment to Cultural Diversity.
Grácio, Maria Cláudia Cabrini	2016	Acoplamento bibliográfico e análise de cocitação: revisão teórico-conceitual
Grácio, Maria Cláudia Cabrini; Oliveira, Ely Francina Tannuri de	2014	Estudos de análise de cocitação de autores: uma abordagem teórico-metodológica para a compreensão de um domínio
Graf, Ann M.	2020	Domain Analysis Applied to Online Graffiti Art Image Galleries to Reveal Knowledge Organization Structures Used Within an Outsider Art Community
Guerrero, Gladys Yolanda Pasuy; Montoya, Claudia Liliana Agudelo	2022	Competencia léxica y procesos de variación semántica en los verbos del parlache
Guimarães, José Augusto Chaves; Matos, Daniela Fernanda de Oliveira; Santos, André Ynada dos; Sales, Rodrigo de Sales	2015	A dimensão conceitual da organização do conhecimento no universo científico da ISKO: uma análise de domínio a partir dos congressos da ISKO-Brasil, ISKO-Espanha, ISKO-América Do Norte e ISKO-França.
Hariri, Nadjla; Kazerani, Maryam; Moghaddam-Kiya, Reza	2013	Evaluation of Persian thesauri of basic sciences based on ANSI / NISO 39.19-2005 standard.
Hartel, Jenna	2010	Managing documents at home for serious leisure: a case study of the hobby of gourmet cooking
Hernandez, Bianca Ferreira; Vital, Luciane Paula	2022	Estudos de gênero para a organização do conhecimento no campo da Arquivologia
Hjørland, Birger	2002a	Domain analysis in information science: Eleven approaches - traditional as well as innovative
Hjørland, Birger	2004b	Arguments for Philosophical Realism in Library and Information Science.
Hjørland, Birger	2005	Empiricism, rationalism and positivism in library and information science
Hjørland, Birger	2009	Concept theory.
Hjørland, Birger	2013b	Citation analysis: A social and dynamic approach to knowledge organization
Hjørland, Birger	2015	Are relations in thesauri 'context-free, definitional, and true in all possible worlds'?
Hjørland, Birger	2017	Domain Analysis.
Hjørland, Birger	2022	Science, Part II: The Study of Science.
Hjørland, Birger	2023	Terminology
Hjørland, Birger; Albrechtsen, Hanne	1995	Toward a New Horizon in Information Science: Domain-Analysis.
Hu, Kai; Wu, Huayi; Qi, Kunlun; Yu, Jingmin; Yang, Siluo; Yu, Tianxing; Zheng, Jie; Liu, Bo	2018	A domain keyword analysis approach extending Term Frequency-Keyword Active Index with Google Word2Vec model

Ibekwe-SanJuan, Fidelia	2006	Constructing and maintaining knowledge organization tools: a symbolic approach.
Krooks, David A.; Lancaster, F. Wilfrid.	1993	The evolution of guidelines for thesaurus construction
Lee, Jae Yun; Kim, Heejung; Kim, Pan Jun	2010	Domain analysis with text mining: Analysis of digital library research trends using profiling methods
León, David Leonardo García	2014	Reflexiones en torno a la situación sociolingüística de las lenguas criollas de base léxica inglesa del Caribe
Lingiardi, Vittorio; Carone, Nicola; Semeraro, Giovanni; Musto, Cataldo; D'Amic, Marilisa; Brena, Silvia	2020	Mapping Twitter hate speech towards social and sexual minorities: a lexicon-based approach to semantic content analysis
López, Miriam Tomás	2004	Diseño de un tesoro conceptual de folklore de la Región de Murcia@es
López-Huertas, María J.	2007	Gender Terminology and Indexing Systems: The Case of Woman's Body, Image and Visualization
Maculan, Benildes Coura; Lima, Gercina Ângela Borém de Oliveira	2014	Relacionamentos em tesouros: o valor semântico dos verbos
Mai, Jens-Erik	2005	Analysis in indexing: document and domain centered approaches
Mai, Jens-Erik	2008	Actors, domains, and constraints in the design and construction of controlled vocabularies
Martines, Alexandre Robson; Almeida, Carlos Cândido	2021	A linguagem documentária e o signo: uma análise sobre termo e conceito
Martines, Alexandre Robson; Moreira, Walter; Almeida, Carlos Cândido	2022	Do signo ao tesouro: contribuições de três correntes da linguagem
Martinez-Gonzalez, M. Mercedes; Alvite-Diez, Maria-Luisa	2020	A semantic web methodological framework to evaluate the support of integrity in thesaurus tools
Martín-Moncunill, David; Sicilia-Urban, Miguel-Ángel; García-Barriocanal, Elena; Sánchez-Alonso, Salvador	2015	Evaluating the degree of domain specificity of terms in large terminologies.
Menard, Elaine; Mas, Sabine; Alberts, Inge	2010	Faceted classification for museum artefacts A methodology to support web site development of large cultural organizations
Miguel, Sandra; Moya-Anegón, Félix; Herrero-Solana, Víctor	2008	New approach to institutional domain analysis: Multilevel research fronts structure
Milani, Suellen Oliveira; Chaves Guimaraes, Jose Augusto	2010	Bias in the indexing languages Theoretical approaches about feminine issues
Moraes, Miriam Gontijo de	2014	Linguagens documentárias e a construção do pensamento crítico
Moraes, Miriam Gontijo de	2018	Tensão identitária e organização do conhecimento
Morando, María Agustina	2021	Los mecanismos de innovación léxica y la reconfiguración de la modernidad indígena en guaraní chaqueño
Moreiro, Jose Antonio	2018	Semantic vocabularies adaptation to the networked digital environment: lexicon, meaning and semantic relationships
Mostafa, Solange Puntel; Cruz, Denise Viuniski da Nova	2010	A importância do empirismo inglês para as linguagens documentárias
Motta, Marta; Tiberi, Melissa	2009	Riflessi dello standard britannico BS 8723 nel Thesaurus del Nuovo soggettarlo.

Moura, Maria Aparecida	2009	Informação, ferramentas ontológicas e redes sociais ad hoc: a interoperabilidade na construção de tesouros e ontologias
Moura, Maria Aparecida; Assis, Juliana	2010	Social network, indexing languages and organization of knowledge: a semiotic approach
Moya-Anegón, Félix; Herrero-Solana, Víctor; Jiménez-Contreras, Evaristo	2006	A connectionist and multivariate approach to science maps: the SOM, clustering and MDS applied to library science research and information
Moya-Anegón, Félix; Vargas-Quesada, Benjamón; Herrero-Solana, Víctor; Chinchilla-Rodríguez, Zaida; Corera-Elvaz, elena; Munoz-Fernandez, Francisco J.	2004	A new technique for building maps of large scientific domains based on the cocitation of classes and categories
Nascimento, Denise Morado; Marteleto, Regina Maria	2008	Social field, domains of knowledge and informational practice
Nielsen, Marianne Lykke	2001	A framework for work task based thesaurus design.
Nunes, Flávia Rodrigues Elias; Maculan, Benildes Coura Moreira Dos Santos; Almeida, Maurício Barcellos	2021	Aporte da linguística no mapeamento dos sentidos dos verbos em tesouros
Oliveira, Dalgiza Andrade; Araújo, Ronaldo Ferreira de	2012	Construção de linguagens documentárias em sistemas de recuperação da informação: a importância da garantia do usuário
Olson, Hope A.	2007	How We Construct Subjects: A Feminist Analysis.
Paiva, Paula Tavares Pinto; Camargo, Diva Cardoso de; Xatara, Cláudia Maria	2008	Uma reflexão sobre a elaboração de um léxico bilíngüe preliminar na subárea de cardiologia a partir do uso de termos encontrados em um corpus paralelo e em dois corpora comparáveis
Pajić, Vesna et al.	2018	Semi-automatic extraction of multiword terms from domain-specific corpora
Peña, Gilberto; Penã, Catalina	2015	Extraction of candidate terms from a corpus of non-specialized, general language
Pinho, Fábio Assis; Guimarães, José Augusto Chaves	2012	Male Homosexuality in Brazilian Indexing Languages: Some Ethical Questions
Pinto, Maria	2008	A user view of the factors affecting quality of thesauri in social science databases
Quesada, Manuel Alejandro Romero; Quintana, Ania Rosa Hernández	2015	O método etnográfico e sua relação com a análise de domínio
Rauch, Mirtha; Barite, Mário	2020	Cultural Warrant Old and New Sights from Knowledge Organization
Reis, Daniela Majorie Akama dos; Fujita, Mariângela Spotti Lopes; Santos, Plácida Leopoldina Ventura Amorim da Costa; Zafalon, Zaira Regina	2018	Tratamento descritivo e temático da informação: recomendações para estudos sobre aspectos semióticos na criação de registros bibliográficos
Reis, Mateus da Silva; Barros, Thiago Henrique Bragato; Santos Junior, Roberto Lopes dos	2021	Proposta metodológica para a elaboração do microtesouro do acervo de medalhística do Museu do Instituto Histórico e Geográfico do Pará
Ribeiro, Ana Rosa Pais; Decourt, Beatriz; Almeida, Tatiana	2017	A representação do domínio? gênero? no âmbito das linguagens documentárias: um mapeamento conceitual em instrumentos terminológicos
Rigo, Sandro José; Alves, Isa Mara da Rosa; Barbosa, Jorge Luis Victória	2015	A Linguistic Approach to Identify the Affective Dimension Expressed in Textual Messages

Rodrigues, Anderson Luiz Cardoso	2012	A Complexidade da cultura amazônica e seu reflexo para a organização e representação da informação
Rodrigues, Anderson Luiz Cardoso	2015	A cultura e a organização do conhecimento: desafios teórico-metodológicos
Sanatjoo, Azam	2010	Development of thesaurus structure through a work-task oriented methodology
Sanchez-Cuadrado, Sonia; Colmenero-Ruiz, Maria-Jesus; Moreiro, Jose-Antonio	2012	Thesauri: Standards and guidelines
Santos, Andréia dos; Costa, Amabile; Barros, Camila Monteiro; Vital, Luciane Paula	2020	Representação terminológica da população negra em tesouros
Sarvabhotla, Kiran; Pingali, Prasad; Varma, Vasudeva	2011	Sentiment classification: a lexical similarity based approach for extracting subjectivity in documents
Schneider, Jesper W.	2005	Naming clusters in visualization studies: Parsing and filtering of noun phrases from citation contexts
Schneider, JW; Borlund, P	2004	Introduction to bibliometrics for construction and maintenance of thesauri - Methodical considerations
Schneider, JW; Borlund, P	2002	Preliminary study of the potentiality of Bibliometric methods for the construction of thesauri
Shiri, Ali Asghar; Revie, Crawford; Chowdhury, Gobinda	2002	Thesaurus-assisted search term selection and query expansion: A review of user-centred studies
Shuter, Janet	2008	New and revised indexing standards.
Silva, Jonathas Luiz Carvalho; Farias, Maria Giovanna Guedes	2013	Reflexões teóricas sobre a construção paradigmática da Ciência da Informação: considerações acerca do(s) paradigma(s) cognitivo(s) e social
Smiraglia, Richard P.	2015	Domain Analysis of Domain Analysis for Knowledge Organization: Observations on an Emergent Methodological Cluster.
Šnajder, Jan; Bašić, B. Dalbelo; Tadić, Marko.	2008	Automatic acquisition of inflectional lexica for morphological normalisation,
Song, Yanhui; Lei, Lixin; Wu, Lijuan; Chen, Shiji	2023	Studying domain structure: a comparative analysis of bibliographic coupling analysis and co-citation analysis considering all authors.
Sousa, Brisa Pozzi de; Tolentino, Vinicius de Souza	2017	Aspectos machistas na organização do conhecimento: a representação da mulher em instrumentos documentários
Souza, Jacqueline A.; Pestana, Olivia	2017	Elos interdisciplinares para estruturação semântica com vistas à organização da informação
Stock, Wolfgang G.	2010	Concepts and semantic relations in information science
Tennis, Joseph T.	2012	Com o que uma análise de domínio se parece no tocante a sua forma, função e gênero?
Thellefsen, Martin Muderspach	2023	Domain analytical information and knowledge organization: investigating the externalist and internalist conception of information
Tiberi, Melissa; Mazzocchi, Fulvio	2007	La gestione della polisemia nei thesauri: il caso dei termini filosofici.
Tomaszuk, Dominik	2022	CoVoMe: New Methodology for Building Controlled Vocabulary
Vital, Luciane Paula; Hernandez, Bianca Ferreira; Santos, Andreia dos	2019	Análise de conteúdo da produção científica da Organização do Conhecimento em sua dimensão política e social no Brasil

Wang, Lin; Qiu, Junping	2021	Domain analytic paradigm: a quarter century exploration of fundamental ideas in information science
Wang, Shengang	2019	The Intellectual Landscape of the Domain of Culture and Ethics in Knowledge Organization: An Analysis of Influential Authors and Works
Weiss, Leila Cristina; Brascher, Marisa	2016	Relações semânticas em tesouros: contribuições da abordagem pragmática
Will, Leonard	2012	The ISO 25964 Data Model for the Structure of an Information Retrieval Thesaurus.
Yoon, JungWon	2009	Towards a user-oriented thesaurus for non-domain-specific image collections
Zeng, Marcia Lei; Chan, Lois Mai	2004	Trends and Issues in Establishing Interoperability Among Knowledge Organization Systems.
Zhao, Dangzhi; Strotmann, Andreas	2022	Intellectual structure of information science 2011-2020: an author co-citation analysis