

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO
DEPARTAMENTO DE CIÊNCIAS DA INFORMAÇÃO
CURSO DE BIBLIOTECONOMIA

WALESKA DA SILVA ROCHA

**SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO: UMA ANÁLISE DE
DOMÍNIO DA NORTH AMERICAN SYMPOSIUM ON KNOWLEDGE
ORGANIZATION (2007-2021)**

Porto Alegre

2024

WALESKA DA SILVA ROCHA

**SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO: UMA ANÁLISE DE
DOMÍNIO DA NORTH AMERICAN SYMPOSIUM ON KNOWLEDGE
ORGANIZATION (2007-2021)**

Trabalho de Conclusão de Curso apresentado como requisito parcial para a obtenção do título de Bacharel em Biblioteconomia, pela Faculdade de Biblioteconomia e Comunicação da Universidade Federal do Rio Grande do Sul.

Orientador: Prof. Dr. Thiago Henrique Bragato Barros

Porto Alegre

2024

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Dr. Carlos André Bulhões Mendes

Vice-reitora: Profa. Dra. Patricia Helena Lucas Pranke

FACULDADE DE BIBLIOTECONOMIA E COMUNICAÇÃO

Diretora: Profa. Dra. Ana Maria Mielniczuk de Moura

Vice-Diretora: Profa. Dra. Vera Regina Schmitz

DEPARTAMENTO DE CIÊNCIAS DA INFORMAÇÃO

Chefe: Prof. Dr. Rene Faustino Gabriel Júnior

Chefe Substituta: Profa. Dra. Caterina Marta Groposo Pavão

COMISSÃO DE GRADUAÇÃO DO CURSO DE BIBLIOTECONOMIA

Coordenadora: Profa. Dra. Maria Lucia Dias

Coordenadora Substituta: Profa. Dra. Helen Rose Flores de Flores

CIP - Catalogação na Publicação

Rocha, Waleska da Silva
Sistemas de Organização do Conhecimento: uma
análise de domínio da North American Symposium on
Knowledge Organization (2007-2021) / Waleska da Silva
Rocha. -- 2024.
72 f.
Orientador: Thiago Henrique Bragato Barros.

Trabalho de conclusão de curso (Graduação) --
Universidade Federal do Rio Grande do Sul, Faculdade
de Biblioteconomia e Comunicação, Curso de
Biblioteconomia, Porto Alegre, BR-RS, 2024.

1. Sistemas de Organização do Conhecimento. 2.
Representação. 3. Análise de domínio. 4. ISKO-C/US. 5.
Sketch Engine. I. Barros, Thiago Henrique Bragato,
orient. II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da UFRGS com os dados fornecidos pelo(a) autor(a).

Faculdade de Biblioteconomia e Comunicação

Rua Ramiro Barcelos, 2705 – Campus Saúde

Bairro Santana, Porto Alegre/RS – CEP 90035-007

Telefone: (51) 3308-5067

E-mail: fabico@ufrgs.br

WALESKA DA SILVA ROCHA

**SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO: UMA ANÁLISE DE
DOMÍNIO DA NORTH AMERICAN SYMPOSIUM ON KNOWLEDGE
ORGANIZATION (2007-2021)**

Trabalho de Conclusão de Curso
apresentado como requisito parcial para a
obtenção do título de Bacharel em
Biblioteconomia, pela Faculdade de
Biblioteconomia e Comunicação da
Universidade Federal do Rio Grande do Sul.

Aprovado em: 07 de fevereiro de 2023.

BANCA EXAMINADORA

Prof. Dr. Thiago Henrique Bragato Barros
Universidade Federal do Rio Grande do Sul (UFRGS)
Orientador

Profa. Dra. Rita do Carmo Ferreira Laipelt
Universidade Federal do Rio Grande do Sul (UFRGS)
Examinadora

Camila Pelizzer
Graduada em Psicologia e Mestranda em Ciência da Informação
Universidade Federal do Rio Grande do Sul (UFRGS)
Examinadora

Ao meu pai e melhor amigo,
Vando Manoel Rocha.

AGRADECIMENTOS

Gostaria de expressar minha gratidão a todos aqueles que me apoiaram ao longo da minha trajetória acadêmica e no desenvolvimento deste Trabalho de Conclusão de Curso. Dedico esse espaço aos familiares, amigos e colegas que eternamente terão minha admiração e respeito. Aprendi que uma graduação não é uma conquista individual, nem apenas uma comprovação de aptidão intelectual. Para crescer em uma área, o companheirismo, confiança e empatia são essenciais.

Meu primeiro agradecimento é dirigido ao meu pai, Vando Manoel Rocha, que, com muito sacrifício, priorizou minha educação e nunca subestimou as minhas capacidades intelectuais. Ensinou-me que, para tudo, nessa vida é imprescindível fé e dedicação. Valores que foram fundamentais no meu desempenho acadêmico nos últimos anos. Obrigada por respeitar e defender minhas decisões pessoais de vida acima das pressões sociais associadas ao meu gênero.

À minha melhor amiga e colega de Biblioteconomia, Ná da Silva Aquino, meu maior suporte emocional nesses últimos quatro anos. Nossa relação vai além dos trabalhos realizados juntas, as nossas risadas e conversas sinceras foram o meu combustível para seguir em frente. Agradeço por me fazer acreditar novamente nas pessoas em todos os momentos em que estava desacreditada. Desejo que a nossa amizade perdure para além da universidade, até o final das nossas vidas.

Também gostaria de agradecer às bibliotecárias e supervisoras de estágio Alexandra Naymayer Corso e Adriana Schwanck Bittencourt que foram essenciais no meu percurso acadêmico e aperfeiçoamento profissional. Levarei comigo lições valiosas sobre competência, respeito e dedicação. A valorização e a confiança depositadas em mim serão para sempre meus motivadores para me tornar no futuro uma profissional cada vez melhor. Dedico também meus eternos agradecimentos à Clarice Heck e à equipe da Academia de Polícia Militar, pelo apoio e carinho nos últimos anos. Guardo com carinho os dois anos da minha graduação que compartilhei com todos.

Agradeço ao meu orientador, Thiago Henrique Bragato Barros, pela orientação e confiança. Desde o segundo semestre, sabia que seguiria minha carreira na pesquisa, mas não tinha ideia de como atingir esse objetivo. Em 2021/1, durante a pandemia, acompanhei a disciplina de “Linguística Aplicada aos Arquivos, Bibliotecas e Museus” ministrada pelo Thiago.

Ao término da última aula daquele semestre, o professor fez um discurso a respeito da importância da atuação dos pesquisadores na Organização do Conhecimento que, na época, despertou meu interesse na área. No ano seguinte, consegui uma bolsa de iniciação científica com o professor Thiago e pude confirmar qual seria minha linha de pesquisa dali em diante. Por isso, aproveito brevemente o momento para fazer esse relato e agradecer a oportunidade que me foi dada.

Por fim, agradeço por todas as experiências que a UFRGS me proporcionou e todas as amizades construídas nesses últimos anos. Não se trata de um final de um ciclo, mas de uma etapa alcançada em um longo caminho a ser percorrido.

Expert, texpert choking smokers,
Don't you think the joker laughs at you?
See how they smile,
Like pigs in a sty, see how they snied.
I'm crying.

Semolina pilchards
Climbing up the Eiffel Tower.
Element'ry penguin singing Hare Krishna,
Man, you should have seen them kicking
Edgar Allan Poe (Lennon; McCartney, 1967).

RESUMO

Dada a diversidade conceitual na área da Organização do Conhecimento, este trabalho visa, através de uma pesquisa de natureza básica e abordagem mista, analisar o capítulo norte-americano da *International Society for Knowledge Organization* (ISKO-C/US). Para isso, propõe analisar temática, semântica e discursivamente os anais publicados pela comunidade discursiva nos eventos da *North American Symposium on Knowledge Organization* (NASKO), no período de 2007 a 2021. A fim de compreender as perspectivas conceituais em relação aos Sistemas de Organização do Conhecimento e à Representação. A metodologia aplicada é a Análise de Domínio, que, a partir de suas abordagens, possibilita o estudo de domínios de conhecimento enquanto comunidades de discurso e pensamento. Neste trabalho, as seguintes abordagens são empregadas: estudos históricos, epistemológicos, estruturais e institucionais e a análise de discurso. O *corpus* da NASKO contempla oito volumes com 132 documentos (127 artigos e 5 resumos expandidos) de tamanho médio (583,678 palavras). Esse conjunto textual é analisado por meio do emprego de enunciados-pivô e do programa de análise textual *Sketch Engine*. Das ferramentas disponibilizadas pelo *software*, utilizaram-se: *Wordlist* (levantamento de frequência das palavras), *Keywords* (levantamento da recorrência semântica das palavras), *Word Sketch* (análise de colocações), *Word Sketch Difference* (análise comparativa entre as colocações de duas palavras) e *Concordance* (para análise do contexto). Os resultados evidenciaram uma ênfase em ferramentas de representação e organização no contexto das bibliotecas. Os Sistemas de Organização do Conhecimento demonstraram ocupar espaço entre os assuntos de interesse, em especial as ontologias. Além de apresentar um tradicionalismo nos tópicos que aborda, ao mesmo tempo que constantemente visa ampliar e inovar suas discussões.

Palavras-chave: Sistemas de Organização do Conhecimento; Representação; Análise de domínio; ISKO-C/US; *Sketch Engine*.

ABSTRACT

Given the conceptual diversity in the field of Knowledge Organization, this work aims to analyze the International Society for Knowledge Organization: Chapter for Canada and the United States (ISKO-C/US) using a basic research and mixed approach. To this end, the study proposes to analyze thematically, semantically, and discursively the proceedings published by the discursive community in the events of the North American Symposium on Knowledge Organization (NASKO) from 2007 to 2021. The analysis aims to understand the theoretical perspectives of NASKO about Knowledge Organization Systems and Representation. The applied methodology is Domain Analysis, which, based on its approaches, enables study of knowledge domains as communities of thought and discourse. The following approaches employed in this work are historical, epistemological, structural and institutional studies and discourse analysis. The NASKO corpus comprises eight volumes with 132 documents (127 articles and five expanded abstracts) of medium size (583,678 words). The analysis of this text collection employed conceptual-terminological terms and the text analysis program Sketch Engine. Among the tools provided by the software, we utilize five: Wordlist (for surveying word frequency), Keywords (for examining semantic recurrence of words), Word Sketch (for analysis of collocations), Word Sketch Difference (for comparative analysis between the co-occurrences of two words), and Concordance (for context analysis). The results highlighted an emphasis on representation and organization tools in the context of libraries. Knowledge Organization Systems were shown to fill the gap between the topics of interest, especially ontologies. In addition, the discursive community presents a traditional approach to the subjects it covers while consistently aiming to broaden and innovate its discussions.

Keywords: Knowledge Organization Systems; Representation; Domain Analysis; ISKO-C/US; Sketch Engine.

LISTA DE SIGLAS E ABREVIATURAS

AAT	<i>Art & Architecture Thesaurus</i>
ANT	Teoria Ator-Rede
ATED	<i>Australian Thesaurus of Education Descriptors</i>
AACR2	Código de Catalogação Anglo-Americano (2ª Edição)
CDD	Classificação Decimal de Dewey
CI	Ciência da Informação
FRBR	Requisitos Funcionais para Registros Bibliográficos
IA	Inteligência Artificial
ISKO	<i>International Society for Knowledge Organization</i>
ISKO-BRASIL	Sociedade Brasileira para Organização do Conhecimento
ISKO-C/US	<i>International Society for Knowledge Organization: Chapter for Canada and United States</i>
LCSH	Cabeçalhos de Assunto da Biblioteca do Congresso
LOD	<i>Linked Open Data</i>
NASKO	<i>North American Society for Knowledge Organization</i>
OBO	<i>Open Biological and Biomedical Ontologies</i>
OC	Organização do Conhecimento
ORC	Organização e Representação do Conhecimento
OWL	<i>Ontology Web Language</i>
RDA	<i>Resource Description and Access</i>
SOC	Sistemas de Organização do Conhecimento
TTI	Tratamento Temático da Informação

SUMÁRIO

1	INTRODUÇÃO.....	13
2	REFERENCIAL TEÓRICO.....	16
2.1	A comunidade discursiva e o discurso.....	16
2.1.1	North American Symposium on Knowledge Organization (NASKO).....	17
2.2	Organização e representação do conhecimento.....	18
2.2.1	Sistemas de Organização do Conhecimento.....	19
3	METODOLOGIA.....	21
3.1	Corpus textual da NASKO e enunciados-pivô.....	21
3.2	Análise de domínio enquanto procedimento.....	23
3.2.1	Sexta abordagem: estudos históricos de estruturas de informação e serviços em domínios.....	25
3.2.2	Oitava abordagem: estudos epistemológicos e críticos de diferentes paradigmas, pressupostos e interesses em domínios.....	25
3.2.3	Décima abordagem: estudos de estruturas e instituições em comunicação científica e profissional em um domínio.....	26
3.2.4	Décima terceira abordagem: Análise de discurso da Análise de domínio.....	26
3.3	Tipo da análise de domínio.....	27
4	SKETCH ENGINE.....	29
4.1	Aspectos linguísticos.....	29
4.1.1	Lexema.....	29
4.1.2	Classe de palavras lexicais.....	30
4.1.3	Colocações.....	31
4.2	Ferramentas.....	32
5	RESULTADOS.....	33
5.1	Wordlist.....	33
5.2	Keywords.....	36

5.2.1	Single-words.....	36
5.2.2	Multi-words terms.....	39
5.3	Word Sketch.....	42
5.4	Word Sketch Difference.....	49
6	CONSIDERAÇÕES FINAIS.....	52
	REFERÊNCIAS.....	55
	APÊNDICE A - LISTA COMPLETA DA WORDLIST.....	62
	APÊNDICE B - LISTA COMPLETA DAS KEYWORDS (SINGLE-WORDS).....	66
	APÊNDICE C - LISTA COMPLETA DAS KEYWORDS (MULTI-WORDS TERMS).....	69

1 INTRODUÇÃO

A organização do conhecimento (OC) é uma área em constante aperfeiçoamento teórico e científico. Guimarães, Matos, Santos e Sales (2015, p. 14) corroboram com essa perspectiva, ao afirmar que este campo está em “[...] fase de consolidação, em busca da construção e delimitação de sua própria base conceitual e seus limites, em virtude de distintas influências teóricas”. Essa questão também é percebida na sociedade científica internacional mais relevante da área, a *International Society for Knowledge Organization* (ISKO).

A ISKO foi fundada na cidade de Frankfurt, Alemanha, em 22 de julho de 1989, pelos membros da *German Society for Classification* (The founding [...], 1989). Seus fundadores incluíram, Ingetraut Dahlberg, Robert Fugmann, Alfred Gerstenkorn, Götz Greiner, Barbara Kelm, Norbert Meder, Padmini Raj, Winfried Schmitz-Esser, Ingeborg Stoltzenburg e Rudolf Ungvary (The founding [...], 1989; Szostak; Ohly, 2020). Hodiernamente, abrange cerca de 600 membros ao redor do mundo, atuantes em diversos campos além da Ciência da Informação (CI), tais como Filosofia, Linguística, Ciência da Computação e Informática Médica (ISKO, 2023b; ISKO-BRASIL, [202-?]). Além disso, os seguintes capítulos regionais integram a comunidade da ISKO: Brasil, Canadá/Estados Unidos (América do Norte), República Popular da China, França, Alemanha/Áustria/Suíça, Índia, Irã, Itália, Países Baixos (Bélgica, Holanda e Luxemburgo), Magrebe, Polônia, Singapura, Espanha/Portugal, Reino Unido, África Ocidental, entre outros (ISKO, 2023a).

Em meio à interdisciplinaridade e expressiva diversidade regional, torna-se difícil atribuir um perfil da ISKO no todo. Guimarães (2014) relembra a preocupação da *International Society for Knowledge Organization* em compreender a sua própria identidade enquanto área do conhecimento. Fato que ocorre devido, segundo o autor, “[...] a uma larga tradição da área, no decorrer de mais de um século, voltada para o estudo e o desenvolvimento de instrumentos e de produtos” (Guimarães, 2014, p. 15). Por isso, estudar e analisar individualmente as comunidades mais representativas dessa sociedade se torna uma via para definir as principais perspectivas e os interesses teóricos dos indivíduos que a compõem.

O estudo em questão se concentra em analisar o capítulo norte-americano da ISKO (ISKO-C/US), composto por Canadá e Estados Unidos. Dado que é uma

unidade que conta com a participação de pesquisadores relevantes na OC e a partir de suas publicações no evento *North American Symposium on Knowledge Organization* (NASKO) promovem suas atividades teórico-científicas. Em outras palavras, estudar as publicações dos seus anais é, conseqüentemente, essencial para compreender o capítulo norte-americano.

Tendo em vista que a OC se desdobra em diversas matérias, define-se como cerne temático, os *Sistemas de Organização do Conhecimento* (SOC) e a *representação no campo da Organização do Conhecimento na América do Norte*. Temas que são recorrentes e representativos para a área da Organização do Conhecimento, sendo discutidos de forma aprofundada ao longo do texto. Diante dessa complexidade, isto é, a diversidade conceitual da ISKO e conseqüentemente do capítulo norte-americano, surge o questionamento “Quais são as concepções conceituais e discursivas da *North American Symposium on Knowledge Organization* (NASKO) em relação à representação e aos Sistemas de Organização do Conhecimento (SOC)?”.

Assim, é proposto, analisar temática, discursiva e semanticamente os anais publicados pela *North American Symposium on Knowledge Organization* (NASKO) entre 2007 e 2021. Com o objetivo de observar e compreender as percepções a respeito da representação e Sistemas de Organização do Conhecimento da comunidade discursiva em questão. A respeito dos objetivos específicos, esses funcionam de forma “[...] intermediária e instrumental, permitindo, de um lado, atingir o objetivo geral e, de outro, aplicá-lo a situações particulares” (Markoni; Lakatos, 2003, p. 219). Os objetivos específicos propostos para atingir o objetivo geral são:

- a) revisar os anais da NASKO publicados no período de 2007 a 2021;
- b) realizar o levantamento métrico dos enunciados-pivô selecionados por meio do programa de análise textual *Sketch Engine*;
- c) identificar os principais fenômenos semânticos e temáticos a partir dos enunciados-pivô estabelecidos;
- d) elaborar uma análise de domínio descritiva através dos levantamentos e características identificadas.

O interesse no tema partiu do desejo pessoal da autora de aprofundar o seu estudo desenvolvido na Iniciação Científica a respeito das atividades discursivas da NASKO. O estudo está atrelado ao projeto de pesquisa intitulado, “Perspectivas teórico-metodológicas sobre Representação e Sistemas de Organização do Conhecimento: uma análise da literatura no âmbito da International Society for Knowledge Organization — ISKO”, financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Sendo o capítulo norte-americano da ISKO um dos objetos de análise que integram o projeto.

Além disso, é possível notar uma contribuição promissora do estudo realizado para a Organização do Conhecimento, uma vez que se dedica a examinar as principais tendências teóricas de uma comunidade influente na área. A análise permite constituir um perfil teórico que ainda é incerto, uma vez que não há trabalhos que tenham se dedicado a esse tipo de investigação. A grande parte dos trabalhos se propõem a uma análise inteiramente bibliométrica e/ou investigam mais de um capítulo da ISKO, dissertando com pouca profundidade a respeito das tendências conceituais do capítulo norte-americano. Partindo desse ponto, a pesquisa se diferencia ao realizar uma análise tanto quantitativa como qualitativa a respeito da ISKO-C/US.

2 REFERENCIAL TEÓRICO

Nesta seção serão apresentados os referenciais teóricos que nortearam o estudo. Dentre os assuntos desenvolvidos estão: o conceito de comunidade discursiva e discurso, o domínio da *North American Symposium on Knowledge Organization*, a Organização e Representação do Conhecimento e os Sistemas de Organização do Conhecimento

2.1 A comunidade discursiva e o discurso

Muitas são as nomenclaturas para um domínio, são algumas dessas: “comunidades de prática”, “comunidades discursivas”, “comunidades epistêmicas” “literatura”, “especialidade” (Tennis, 2003; Smiraglia, 2012; Hjørland, 2017), entre outras. Nenhum desses termos é considerado incorreto, contudo as particularidades do estudo definem quais, entre as muitas definições, poderão ser aplicáveis. Vale salientar, porém, que o domínio não é uma extensão delimitada, e sim um campo que passa constantemente por mutações (Hjørland, 2017). Dado que o foco principal está ligado à discursividade do domínio estudado, define-se que “comunidades discursivas” corresponde aos interesses principais da pesquisa.

O discurso é definido por Pêcheux, conforme Orlandi (2005, p. 11), “[...] como sendo efeito de sentidos entre locutores, um objeto sócio-histórico em que o [linguístico] está pressuposto”. A presença da palavra *locutores* no plural indica a importância do processo comunicativo, aliado aos processos discursivos. Conforme Pêcheux (1995, p. 91), “[...] o sistema da língua é, de fato, o mesmo para o materialista e para o idealista, para o revolucionário e para o reacionário, para aquele que dispõe de um conhecimento dado e para aquele que não dispõe desse conhecimento”. Contudo, os processos discursivos de cada um desses agentes podem divergir, mesmo que a língua seja a base comum desses processos. Pois, “[...] a discursividade não é a fala (*parole*¹), isto é, uma maneira individual “concreta” de habitar a “abstração” da língua; não se trata de um uso, de uma utilização ou da realização de uma função” (Pêcheux, 1995, p. 91). Trata-se, então, de um construto

¹ Cf. SAUSSURE, 2006, p. 21.

social, não puramente cognitivo, que proporciona um objeto passível de ser analisado.

Por isso, o objeto estudado são os anais da NASKO que por sua vez estão organizados em um conjunto textual, um *corpus*². De acordo com Sinclair (1991, p. 102), fatores como “[...] criatividade, inconsciência, conveniência, desatenção, confusão e a necessidade de expressar o comum” influenciam os desvios normativos em um texto. Para isso, são utilizados *softwares* que analisam o que é mais ou menos comum, o que possibilita o *isolamento do sistema abstrato* (Sinclair, 1991).

A partir da perspectiva Pêcheuxiana, tem-se um contraste entre a linguagem e a discursividade, sendo a primeira aberta a ambiguidades e a segunda amparada na materialidade histórica (Orlandi, 2005). O que valida a análise de um *corpus* através do seu discurso, ao demonstrar que existe uma materialidade a ser investigada. Uma materialidade que surge de componentes como intencionalidade e semântica.

2.1.1 North American Symposium on Knowledge Organization (NASKO)

A ISKO-C/US começou em 2007, com o primeiro evento da *North American Symposium on Knowledge Organization*. Somente se estabelecendo oficialmente como capítulo regional norte-americano da ISKO em 2009, conforme o estatuto local aprovado pela NASKO e o estatuto para capítulos regionais redigido pela *International Society for Knowledge Organization* (ISKO-C/US, c2009-2010a).

De acordo com o *ISKO News 13*, publicado no periódico *Knowledge Organization* (v.20, n.2) em 1993, “Um Capítulo Regional consiste em um grupo de membros da ISKO vivendo em uma região geográfica bem definida” (ISKO [...], 1993, p. 91). O propósito da ISKO-C/US, assim como de outros capítulos regionais, é atender aos interesses da ISKO e possibilitar que os membros sejam atuantes em suas respectivas regiões (ISKO [...], 1993).

O capítulo norte-americano (composto por Canadá e Estados Unidos) segue em atividade até o momento. Entre 2007 a 2021, os temas abordados pela NASKO foram (ISKO-C/US, c2009-2010b):

² “No vocabulário da ciência, **corpus** designa uma grande coleção e às vezes exaustiva, de documentos ou dados [...]” (CHARAUDEAU; MAINGUENEAU, 2002, p. 148, tradução nossa).

- a) **2007** — *Knowledge Organization*;
- b) **2009** — *Pioneering North American Contributions to Knowledge Organization*;
- c) **2011** — *Expanding Our Horizons, Evaluating Our Parameters*;
- d) **2013** — *Transition Cultures, Transition KO: Evolving Exploration, Critical Reflection, and Practical Work*;
- e) **2015** — *Producing Knowledge Organization*;
- f) **2019** — *Community and Computation*;
- g) **2021** — *Resilience, Resistance, and Reflection: Knowledge Organization at a Crossroads*.

Tendo em vista as últimas temáticas abordadas pela comunidade nos últimos anos, percebe-se que a preocupação da ISKO em compreender sua própria base epistemológica e conceitual, também reflete nos interesses do capítulo norte-americano. Dado que os temas têm como foco questões relacionadas à produção, ao desenvolvimento e à contribuição da comunidade discursiva para a área da Organização do Conhecimento. A última conferência realizada pela comunidade em 2023, intitulada “Who Are We?: Reflecting, Shaping, and Challenging Identity Through Knowledge Organization³” (“Quem nós somos?: Refletindo, Moldando e Desafiando a Identidade através da Organização do Conhecimento”, em português) (ISKO-C/US, c2009-2010b), reforça essa característica da busca pela autocompreensão.

2.2 Organização e representação do conhecimento

A Organização e Representação do Conhecimento (ORC) é uma área que pertence à Ciência da Informação (CI), se dedicando ao conteúdo e sua representação, possibilitando a discussão teórica e o desenvolvimento de metodologias, a fim de proporcionar o acesso à informação (Barros; Laipelt, 2021; Barros, 2021). A partir do seu desenvolvimento, diversas técnicas, atividades e ferramentas podem ser elaboradas: “[...] (1) processos de ORC, indexação, catalogação, análise de assunto, classificação; (2) sistemas de Organização e

³ É importante frisar que o último evento realizado em 2023 não está incluso nessa análise, pois os documentos não estão disponíveis para consulta no site oficial da ISKO-C/US.

Representação do Conhecimento, gerados para esses processos” (Hjørland, 2008; Barros; Laipelt, 2021, p. 441).

São pertinentes para esse campo, estudos a respeito da “[...] construção [de] tesouros, vocabulários controlados, taxonomias, ontologias, ou seja, uma gama de instrumentos-processos que visam à busca e apropriação do conhecimento, tradicionalmente ligada à Ciência da Informação” (Barros, 2021, p. 63). Todos esses processos e sistemas ligados à organização e representação do conhecimento são considerados definidores da OC em seu sentido estrito, de acordo com Hjørland (2008). No que tange ao seu sentido amplo, o conhecimento é organizado na realidade e socialmente (Hjørland, 2008). Isto é, a organização e representação do conhecimento se estendem para além do setor da CI. Encontra-se presente em todos os setores da sociedade que a utilizarem.

2.2.1 Sistemas de Organização do Conhecimento

Conforme Hodge (2000, p. 2, tradução nossa), os SOC possibilitam englobar “[...] todos os tipos de esquemas de organização de informação e promover a gestão do conhecimento”. Para Brascher e Café (2008, p. 8), “[...] são sistemas conceituais que representam determinado domínio por meio da sistematização dos conceitos e das relações semânticas que se estabelecem entre eles”. Hjørland (2015, p. 114, tradução nossa), por sua vez, declara, “De relevância específica para a OC é a afirmação de que uma teoria é um sistema de organização do conhecimento (SOC) e vice-versa: qualquer SOC é, se não uma teoria, pelo menos teórica e ideologicamente carregado”.

Sabe-se, por meio de todas essas percepções, que os Sistemas de Organização do Conhecimento são sistemas que possibilitam organizar e representar um determinado domínio do conhecimento. A depender de sua função e do grupo social que se atinge, a sua função e fundamentação tecnológica tornam-se variáveis (Barros; Laipelt, 2021). Como a pesquisa de Barros, Bastos e Santos (2022), que concluiu que o tesouro tradicional não seria tão adequado quanto a ontologia para o contexto e demandas de um arquivo.

Ou seja, não há um modelo único de SOC, por existirem variadas formas de organização do conhecimento (Hodge, 2000). Dependerá da necessidade e do

contexto a qual se aplica. Os Sistemas de Organização do Conhecimento, conforme Zeng (2008), poderão apresentar um maior ou menor grau de complexidade a depender das funções que fornecem.

De acordo com Hjørland (2015), os SOC irão representar diferentes ênfases em suas relações semânticas. O autor correlaciona esses sistemas com a “escada semântica”, sendo a ontologia o sistema mais rico semanticamente. Os sistemas de organização do conhecimento são as ferramentas que viabilizarão a organização e representação do conhecimento. Por isso, se caracterizam como tema central deste estudo.

3 METODOLOGIA

Trata-se de uma pesquisa de natureza básica que utiliza a abordagem mista, isto é, um estudo qualitativo que parte da análise descritiva e interpretativa de uma comunidade científica; e quantitativa, que utiliza dados métricos fornecidos pelo *software Sketch Engine*. McKim (2017) aponta que tanto a abordagem quantitativa como a qualitativa recebem críticas, quando aplicadas isoladamente. O ponto negativo central destacado de cada uma das abordagens são, respectivamente, o pouco aprofundamento interpretativo e a falta de objetividade. Por isso, como explica McKim (2017), os pesquisadores utilizam o método misto para antever essas problemáticas e resolver as controvérsias entre ambas.

O grupo estudado é o capítulo norte-americano da ISKO, por meio das suas publicações nos anais da NASKO. Seu caráter é descritivo, dado que o estudo identifica as atividades discursivas de um determinado grupo a partir da aplicação da análise de domínio. Isso significa descrever uma população, fenômeno(s) ou estabelecer relações entre variáveis, se utilizando de uma técnica de coleta de dados padronizada (Gil, 2008). Para tal, o procedimento utilizado é a análise de domínio em conjunto com o *software Sketch Engine*.

3.1 Corpus textual da NASKO e enunciados-pivô

Na etapa preliminar, realizou-se uma seleção prévia do 11 enunciados-pivô⁴. Dessa forma, o estudo concentrou seus interesses nas temáticas centrais, a representação e os SOC. Os termos foram extraídos tendo por base o “Diccionario de organización del conocimiento: clasificación, indización, terminología” de Mario Barité (2015) e o artigo de Barros e Laipelt (2021, p.445). Os enunciados-pivô selecionados foram organizados com suas respectivas definições no quadro 1.

⁴ Proveniente do campo linguístico, trata-se de uma junção dos termos “enunciado” (conjunto de orações de tamanhos variáveis) e “pivô”, do método de termos-pivô francês (apanhado supostamente representativo de palavras-chave em um discurso para posterior análise) (CHARAUDEAU; MAINGUENEAU, 2002).

Quadro 1 - Enunciados-pivô selecionados para a análise

ENUNCIADOS-PIVÔ	TRADUÇÃO	DEFINIÇÃO
<i>Information Organization</i>	Organização da Informação	“A organização da informação é, portanto, um processo que envolve a descrição física e de conteúdo dos objetos informacionais” (Brascher; Café, 2008, p. 5).
<i>Information Representation</i>	Representação da Informação	“O produto desse processo descritivo [da organização da informação] é a representação da informação, entendida como um conjunto de elementos descritivos que representam os atributos de um objeto informacional específico” (Brascher; Café, 2008, p. 5).
<i>Knowledge Organization</i>	Organização do Conhecimento	“[...] estuda as leis, os princípios e procedimentos pelos quais se estrutura o conhecimento especializado em qualquer disciplina, com a finalidade de representar e recuperar tematicamente informação contida em documentos de qualquer natureza, por meios eficientes que proporcionem respostas rápidas as necessidades dos usuários” (Barité, 2015, p. 120, tradução própria).
<i>Knowledge Representation</i>	Representação do Conhecimento	“[...] compreende um conjunto de processos de simbolização notacional ou conceitual do saber humano no âmbito de qualquer disciplina. [...] incluem a classificação, a indexação e o conjunto de aspectos informáticos e linguísticos relacionados com a tradução simbólica do conhecimento” (Barité, 2015, p. 136, tradução nossa).
<i>Knowledge Organization and Representation</i>	Organização e Representação do Conhecimento	“[...] é um palco de estudo institucionalizado socialmente e cientificamente dentro da área de Ciência da Informação para discutir teorias e metodologias relacionadas aos vários processos de representação e organização” (Barros; Laipelt, 2021, p. 439-440).
<i>Indexing</i>	Indexação	“[...] diz respeito à identificação do conteúdo do documento por meio do processo de análise de assunto e à representação desse conteúdo por meio de conceitos” (Rubi, 2009, p. 81).
<i>Documentary Representation</i>	Representação documentária	Conforme Fujita (2013), integra duas etapas no processo de indexação: a análise de assunto (ao identificar e selecionar conceitos) e a tradução (ao traduzir os termos para uma linguagem documental).

<i>Knowledge Organization Systems</i>	Sistemas de Organização do Conhecimento	De acordo com Hodge (2000, p. 1, tradução nossa), “[...] abrange todos os tipos de esquemas para organização da informação e promoção da gestão do conhecimento”. Incluem tanto estruturas de menor como de maior complexidade (Hodge, 2000; Zeng, 2008).
<i>Taxonomy</i>	Taxonomia	“[...] divisão de itens em grupos ou categorias ordenados em características particulares” (Zeng, 2008, p. 161, tradução nossa).
<i>Thesaurus</i>	Tesouro	“[...] conjuntos de termos que representam conceitos e a hierarquia, equivalência e relações associativas entre si” (Zeng, 2008, p. 161, tradução nossa).
<i>Ontology</i>	Ontologia	“[...] modelos conceituais específicos que representam relações complexas entre objetos, incluindo as regras e axiomas que faltam em redes semânticas” (Zeng, 2008, p. 162, tradução nossa).

Fonte: Rocha (2023).

Além disso, todos os anais foram previamente reunidos, constituindo um total de 132 documentos (127 artigos e cinco resumos expandidos). Os documentos foram levantados e analisados no período de 26 de setembro de 2022 a 8 de janeiro de 2023, o que resultou em um *corpus* de 583,678 palavras em inglês. A coleta de dados antecipada se deve ao fato de a autora ter feito previamente uma pesquisa a respeito do assunto para a sua bolsa de Iniciação Científica. Além de que essa etapa despense muito tempo, o que não seria viável para ser realizado ao longo do desenvolvimento do trabalho. Independentemente disso, o propósito foi aplicar uma revisão no *corpus* e expandir a análise a fim de realizar uma investigação com maior cobertura.

3.2 Análise de domínio enquanto procedimento

A metodologia aplicada foi a Análise de Domínio desenvolvida em 1995 por Birger Hjørland e Hanne Albrechtsen, que, desde então, é revista e aprimorada por Hjørland (2002, 2004, 2017). Ao defender a necessidade de teorias que estudassem a própria CI, Hjørland (2004, p. 17, tradução nossa) argumenta que “Um campo que não consegue confrontar racionalmente os seus problemas, está em crise”. Como

resolução, Hjørland e Albrechtsen (1995, p. 400, tradução nossa) propuseram a teoria da Análise de Domínio que “[...] estuda os domínios do conhecimento como pensamento ou comunidades de discurso, que são parte da divisão de trabalho da sociedade”. Desde então, tem sido adotada como abordagem metodológica pela área da Organização do Conhecimento.

Essas comunidades discursivas não se tratam de “mentes universais” (Hjørland, 2017), em que não é possível delimitar a sua identidade. Pelo contrário, tendo em consideração que a informação em seu estado intangível precisa ser representada para atingir um estado de tangibilidade (Buckland, 1991), entidades científicas conseguem transcrever com êxito o seu conhecimento adquirido para um meio formal de comunicação científica. Através desses registros (os anais da NASKO, neste caso), têm-se uma matéria de estudo alcançável que reflete os posicionamentos discursivos de um determinado grupo.

A metodologia em questão propõe onze abordagens para estudos de domínio que conforme Hjørland (2017, p. 436-437), são:

1. Produção e avaliação de guias de literatura e gateways de assunto;
2. Produção e avaliação de classificações especiais e tesouros;
3. Pesquisa sobre competências em indexação e recuperação de informações em especialidades;
4. Conhecimento de estudos empíricos com usuários em áreas temáticas;
5. Produção e interpretação de estudos bibliométricos;
6. Estudos históricos de estruturas de informação e serviços em domínios;
7. Estudos de documentos e gêneros em domínios do conhecimento;
8. Estudos epistemológicos e críticos de diferentes paradigmas, pressupostos e interesses em domínios;
9. Conhecimento de estudos terminológicos, LSP (linguagens para fins especiais) e análise de discurso em campos do conhecimento;
10. Estudos de estruturas e instituições em comunicação científica e profissional em um domínio;
11. Conhecimento de métodos e resultados de estudos analíticos de domínio sobre cognição profissional, representação de conhecimento em ciência da computação e inteligência artificial (Hjørland, 2017, tradução nossa).

Dentre as abordagens mencionadas anteriormente, três são foco do estudo: 6 — *Estudos históricos de estruturas de informação e serviços em domínios*; 8 — *Estudos epistemológicos e críticos de diferentes paradigmas, pressupostos e interesses em domínios*; e 10 — *Estudos de estruturas e instituições em comunicação científica e profissional em um domínio*. E a 13 — *Análise de Discurso da Análise de Domínio*, proposta por Smiraglia (2015) e fundamentada por Barros

(2023), em seu artigo “Michael Pêcheux’s Discourse Analysis: An Approach to Domain Analyses”.

3.2.1 Sexta abordagem: estudos históricos de estruturas de informação e serviços em domínios

Estudos históricos são essenciais, à medida que viabilizam a compreensão das origens dos domínios e suas respectivas vertentes. De acordo com Hjørland (2002, p. 436, tradução nossa) é uma abordagem que enriquece o estudo, ao aprofundar e analisar com maior coerência “[...] documentos, organizações, sistemas, conhecimento e informação, perspectiva histórica e métodos históricos [...]”. Segundo Almeida e Dias (2019, p. 31), “[...] organizam tradições, paradigmas, documentos e formas de expressão e suas influências mútuas”.

A partir disso, obtém-se um embasamento que contribui significativamente para análise, tanto do que está de acordo quanto do que diverge com a perspectiva histórica pré-estabelecida daquela organização. O conhecimento prévio desse fator que envolve aspectos históricos é fundamental para a implementação de uma análise de domínio bem fundamentada e interage com a abordagem apresentada a seguir.

3.2.2 Oitava abordagem: estudos epistemológicos e críticos de diferentes paradigmas, pressupostos e interesses em domínios

Comumente atrelados aos estudos históricos, os estudos epistemológicos investigam os pressupostos que estão explícitos ou implícitos por trás das tradições de pesquisa (Hjørland, 2002). De acordo com Smiraglia (2012, p. 111, tradução nossa), “Na Organização do Conhecimento, a atividade chamada análise de domínio é o ato de definir a base de conhecimento conceitual de uma comunidade”, isto é, analisar um domínio tendo por base os pressupostos que norteiam as atividades de uma comunidade discursiva.

Pode-se concluir então, a partir das perspectivas de Hjørland e Smiraglia, que os estudos epistemológicos são indispensáveis para aplicação da análise de domínio. Tendo em vista que além de observar as ocorrências (termos, assuntos,

autores, etc.), o intuito é responder aos “porquês” que se relacionam às suas atividades. A não aplicação incorrerá na superficialidade do estudo (Hjørland, 2002).

3.2.3 Décima abordagem: estudos de estruturas e instituições em comunicação científica e profissional em um domínio

São estudos que podem ser vistos como modeladores de “[...] sistema de instituições, serviços e fontes de informação primária, secundária e terciária, intermediados entre produtores de conhecimento e usuários do conhecimento” (Almeida; Dias, 2019, p. 30). Para compreensão da função de determinados tipos de documentos e serviços de informação é necessário estudar a estrutura organizacional de divisão de trabalho de um domínio (Hjørland, 2002). No caso do estudo desenvolvido, compreender do que se trata a comunidade da NASKO e qual o caráter dos documentos publicados em seus anais.

De acordo com Almeida e Dias (2019), possuem como principal função organizar os atores e instituições centrais. A partir disso, sabe-se que o capítulo norte-americano é organizado por pesquisadores (uma comunidade científica) e seus documentos se encontram em um plano formal de apresentação devido ao seu caráter teórico-científico.

3.2.4 Décima terceira abordagem: Análise de Discurso da Análise de Domínio

De 1995 adiante, houve revisões e colaborações de teóricos da área para o desenvolvimento contínuo da abordagem metodológica. Dentre essas contribuições, integram o corpo do trabalho a décima terceira abordagem proposta por Smiraglia (2015) e, posteriormente, fundamentada por Barros (2023), a Análise de Discurso. De acordo com Pêcheux (1997), se trata de um processo de análise constituído de descrição e interpretação, onde a interpretação deve considerar o contexto em que o enunciado está inserido.

Segundo Smiraglia (2015), uma análise realizada por uma teoria crítica possibilita verificar influências sociais e culturais, contribuindo significativamente para a compreensão de um domínio. Essa criticidade é viabilizada pela bagagem ideológica e teórica daquele que descreve e interpreta o discurso. Partindo da

premissa inicial de que todo enunciado reserva uma intenção, sabe-se então que todo interlocutor, segundo Azeredo (2012), se torna “cúmplice” das variáveis⁵ que compõem o discurso. Essa cumplicidade pode vir a partir de um discurso implícito — proveniente da aptidão interpretativa do interlocutor — e/ou discurso explícito — conjunto de enunciados diretamente expressos pelo locutor (Azeredo, 2012).

Considerando que dentre as influências da análise de discurso subsiste o materialismo histórico (e a linguística) (Orlandi, 2005; Barros, 2021, 2023), sabe-se que o cerne da pesquisa se centra na “dimensão histórica dos processos sociais” sob o discurso que será analisado (Gil, p. 22, 2008).

Para a análise do discurso, a produção textual funciona como um objeto de inscrição daquele e da ideologia; o discurso é um espaço aberto, que parte da língua, já que **a materialidade é o texto e é atravessado pela ideologia, circunscrito por sua própria historicidade**, isto é, **o discurso é, do ponto de vista de sua formulação, um acontecimento e uma estrutura**. (Barros, 2021, p. 69, grifo nosso).

Por essas características, a análise de discurso é uma das abordagens selecionadas, visto que pode contribuir para o desenvolvimento de sistemas de organização do conhecimento e práticas discursivas próprias da organização do conhecimento (Barros, 2023).

3.3 Tipo da análise de domínio

Tennis (2012, p. 9) apresenta dois tipos de análise de domínio que contribuem para a definição de uma finalidade: a “Análise de Domínio Descritiva” que atende ao princípio de uma descrição voltada a pesquisa básica e a “Análise de Domínio Instrumental” que tem por finalidade o desenvolvimento ou revisão de um sistema de informação. Conforme Smiraglia (2012), a metodologia em questão é o coração da OC, porque sem sua aplicação não haveria matéria ontológica para o desenvolvimento de Sistemas de Organização do Conhecimento.

Com base nisso, sabe-se que o estudo aplica a análise de domínio de maneira exclusivamente descritiva, visto que está se investigando e detalhando as atividades discursivas da comunidade da NASKO. Não há pretensão de desenvolver

⁵ As variáveis, de acordo com Azeredo (2012, p. 128-129), se tratam de: “[...] o código linguístico, o interlocutor, o tempo, o espaço, a situação social, o conteúdo, crenças e valores culturais, o texto em processo, outros textos”.

ou avaliar um sistema de informação. O que não anula a possibilidade do resultado de uma proposta se transformar em outra, entretanto, para fins metodológicos é necessário explicitar objetivamente o intuito principal do trabalho (Tennis, 2012).

4 SKETCH ENGINE

O *Sketch Engine* é um *software* utilizado para análise linguística de textos. Trata-se de uma ferramenta que explora o que é recorrente (raro, atípico e de uso emergente) na linguagem de um *corpus* (Sketch Engine, [202-?]c). Através do seu uso é possível organizar, visualizar e delinear as tendências linguísticas e conceituais de uma comunidade discursiva, pois o recurso elabora gráficos e quadros que exemplificam visualmente as tendências na produção intelectual.

Não obstante, coube a aplicação de uma limpeza de dados para manter o rigor e máxima fidelidade aos propósitos do trabalho. Dentre as ferramentas fornecidas pelo *Sketch Engine*, foram pré-selecionadas quatro destas para desenvolvimento, organização gráfica e tratamento analítico dos termos dentro do *corpus* e uma, em especial, foi utilizada como ferramenta auxiliar das demais (totalizando cinco ferramentas descritas posteriormente). O tratamento dos dados foi realizado a partir da seleção dos enunciados-pivô preestabelecidos e das palavras que estão consoantes a área da organização do conhecimento.

4.1 Aspectos linguísticos

Dado que alguns termos linguísticos são utilizados nas etapas descritivas, é essencial defini-los para assegurar a compreensão da análise e elucidar algumas decisões na aplicação do programa. Ao longo do processo de tomada de decisão e limpeza de dados, percebeu-se que não é possível realizar uma análise textual sem ter um conhecimento linguístico prévio. As seções seguintes são dedicadas para exploração desses aspectos linguísticos.

4.1.1 Lexema

Para algumas ferramentas no *Sketch Engine* é necessário estabelecer um tipo de unidade gramatical para ser o objeto de análise. O que é importante analisar? Verbos? Substantivos? Preposições? Artigos? Lexemas? Para os fins desta pesquisa, é certo estabelecer o lexema enquanto critério principal. Para justificar a escolha, é pertinente compreender o que é essa unidade.

Booij (2005) define o lexema (“lemma”, em inglês) como uma noção abstrata da palavra que origina palavras de sentido concreto, as *word forms* (“formas da palavra”). Por exemplo, a partir do lexema *cuidar* (verbo na sua forma infinitiva) derivam suas flexões concretas, *cuida*, *cuidas*, *cuidaram*, *cuidam*, *cuidei*, etc. Sendo assim, um lexema traz dentro de si todas as possíveis variações que nela podem ocorrer. Para este estudo, interessa analisar todas as variações da palavra, por ser pertinente para a proposta do estudo o sentido da palavra na sua totalidade. Uma vez que expande o campo de análise dos enunciados-pivô, possibilitando assim, que a pesquisa não fique engessada a termos específicos. Além disso, as demais ferramentas estão previamente configuradas para analisarem os termos enquanto *lemmas*. O critério de padronização e conformidade é mais uma questão que reforça a decisão.

4.1.2 Classe de palavras lexicais

Após estabelecer o lexema enquanto unidade gramatical da análise, deve-se pensar com quais entes (classes de palavras) esse objeto pode se relacionar e quais dessas relações⁶ interessam à análise. Algumas ferramentas como *Word Sketch* e *Word Sketch Difference* dispõem de um leque de classes gramaticais. Os tipos que serão selecionados variam conforme os propósitos da pesquisa e o comportamento gramatical da língua. Logo, compreender os seus tipos é indispensável para dominar a análise. De acordo com Rijkhoff (2007), as classes de palavras dividem-se em duas formas:

- a) **classe de palavras gramaticais**, que apresentam pouco ou nenhum sentido expresso. Comumente conhecidas como “function words” ou “empty words” (“palavras de função” ou “palavras vazias”, em português);
- b) e **classe de palavras lexicais**, que possuem uma definição semântica estabelecida e podem mudar de significado dependendo do contexto que estão inseridas. Conhecidas como “content words”, isto é, palavras de

⁶ De forma implícita, trata-se das relações colocacionais que serão discutidas posteriormente no decorrer do trabalho.

conteúdo. Na língua inglesa existem quatro classes lexicais principais: verbos, substantivos, adjetivos e advérbios.

A apresentação da definição dessas classes permite estabelecer quais critérios que balizam a análise das colocações no *Sketch Engine*. Uma vez que para esse trabalho o foco principal está nas unidades lexicais que estabelecem uma relação semântica entre si (classes de palavras lexicais). Contudo, é importante salientar que o programa utiliza uma classe denominada “nouns modifiers” (que pode ser traduzida no português para, “modificadores de substantivos”). Modificadores de substantivos atuam principalmente como complementos inseparáveis ou adjuntos de um *head noun* (“substantivo principal”) (Rosenbach, 2007)⁷.

Por fim, os lexemas serão analisados em relação aos verbos e *nouns modifiers*. Pois, Rijkhoff (2007, p. 711, tradução nossa) explica que os verbos são “[...] tipicamente associados com ações ou situações (isto é, entidades temporais) e os substantivos com coisas (entidades espaciais) [...]”.

4.1.3 Colocações

Quanto às colocações, esse é um termo também relacionado às ferramentas *Word Sketch* e *Word Sketch Difference*. Conforme Evert (2009), através de uma perspectiva empírica, colocações dizem respeito à combinação entre termos que costumam estabelecer conexões recorrentes em uma linguagem natural. Um exemplo mencionado pela autora é a relação entre “vaca” e “leite”, não é incomum estabelecer cotidianamente a relação entre essas palavras. No entanto, para evitar que a análise permaneça restrita ao campo “intuitivo”, são empregadas pontuações associativas que, de acordo com Proisl (2019), constituem uma *abordagem orientada estatisticamente*. Assim sendo, no caso do *Sketch Engine*, por mais que existam preconceções, elas são avaliadas por meio do programa tanto para corroboração como para contraposição.

⁷ Rosenbach (2007) explica detalhadamente a diferença entre *nouns modifiers*, *adjectives* e *s-genitives* nas páginas 145-151.

4.2 Ferramentas

Cinco ferramentas do programa foram utilizadas. Conforme *Sketch Engine* ([202-?]b), são essas:

- a) **Wordlist** (“*lista de palavras*”): levantamento da frequência métrica das palavras, ou seja, a quantidade de vezes que cada lexema aparece no *corpus*;
- b) **Keywords** (“*palavras-chave*”): extração da recorrência semântica das palavras que definem o tópico central do *corpus*. Divide-se em *single-words* (“*palavras-únicas*”) que realiza o levantamento de unidades lexicais constituídas de uma palavra e *multi-word terms* (“*termos com múltiplas palavras*”) que se concentra em palavras que formam uma expressão;
- c) **Word Sketch** (“*esboço de palavras*”): ferramenta para identificação de colocações mais recorrentes de um lexema. Isto é, as relações mais recorrentes estabelecidas de uma palavra com as outras. Por meio do seu uso é possível verificar o comportamento gramatical e colocacional;
- d) **Word Sketch Difference** (“*diferenças entre esboços de palavras*”): utilizada para estabelecer uma relação comparativa entre dois esboços de palavras. Dois lexemas são cotejados para examinar, de forma comparada, como funcionam as suas relações gramaticais e colocacionais;
- e) **Concordance** (“*Concordância*”): ferramenta utilizada em paralelo com as demais anteriormente mencionadas. Seu uso permite uma verificação minuciosa do sentido de cada lexema no texto. Auxiliando nas tomadas de decisão e limpeza dos dados obtidos.

5 RESULTADOS

Nesta seção, serão analisados, qualitativa e quantitativamente, os gráficos e as tabelas geradas pelo *Sketch Engine*, com o objetivo de detalhar a organização e considerações alcançadas. No total, 132 documentos (127 artigos e 5 resumos expandidos) foram examinados, o que incluiu todas as publicações da *North American Symposium on Knowledge Organization* entre 2007 e 2021. O *corpus* resultou em objeto de estudo com 583,678 palavras. Sardinha (2003, p. 117-118) afirma que este é um *corpus* de tamanho médio, uma vez que a média⁸ é constituída por 500,000 palavras.

As tabelas apresentam um recorte de 20 dados por lista, pois essa quantia é considerada suficiente para representação e análise terminológica do *corpus*. Além disso, descobriu-se que quanto menor o *score* do termo, menor a sua presença em mais de um documento. Assim sendo, o foco foi a ocorrência dos termos no conjunto textual, com a análise direcionada aos termos de pontuação elevada. Para verificar as listas na íntegra, ver APÊNDICE A — LISTA COMPLETA DA WORDLIST, APÊNDICE B — LISTA COMPLETA DAS KEYWORDS (SINGLE-WORDS) e APÊNDICE C — LISTA COMPLETA DAS KEYWORDS (MULTI-WORDS TERMS).

5.1 Wordlist

A ferramenta tem como função principal viabilizar a análise métrica de um *corpus*. As frequências absolutas das palavras foram analisadas, isto é, o número total de ocorrências dos termos no conjunto textual. Dessa forma, foi possível traçar um perfil terminológico dos anais da NASKO por ordem de frequência⁹. Ao analisar as palavras na *Wordlist*, é indispensável selecionar uma classe de palavras. Dentre as classes disponíveis para análise, o lexema foi o critério adotado.

Após a seleção da classe, a ferramenta gerou uma lista de palavras que passou por uma limpeza de dados. Sinais gráficos, *empty words* (preposição, artigo, conjunções, etc), números, nomes, sobrenomes e abreviaturas foram

⁸ O estudo realizado por Sardinha (2003) teve como base os *corpora* apresentados em conferências de Linguística de *Corpus* entre os anos de 1995 a 1998. A média é obtida por meio da mediana, que é considerada a mais adequada pelo autor, pois representa o valor central da distribuição.

⁹ De acordo Sinclair (2003), esse tipo de análise tem validade principalmente para um *corpus* com mil ou mais palavras.

desconsiderados. Privilegiou-se os dados que possuem alguma relação com o tema estudado. Para obtenção de um parâmetro terminológico de todo o *corpus*, é crucial iniciar a análise utilizando esta ferramenta.

As três primeiras palavras, *knowledge* (“conhecimento”, 3,701 ocorrências), *classification* (“classificação”, 3,398 ocorrências) e *information* (“informação”, 2,587 ocorrências), revelaram-se bastante características da terminologia da OC, uma vez que apresentaram uma consistência temática no escopo estudado. Na quarta posição, percebeu-se uma ênfase nas bibliotecas (“library”, 2,518 ocorrências). Os arquivos não aparecem na tabela 1.

Tabela 1 - Lista de palavras do *corpus* da NASKO (2007-2021)

Ordem	Palavra	Tradução	Frequência absoluta
1º	<i>knowledge</i>	conhecimento	3,701
2º	<i>classification</i>	classificação	3,398
3º	<i>information</i>	informação	2,587
4º	<i>library</i>	biblioteca	2,518
5º	<i>organization</i>	organização	2,472
6º	<i>subject</i>	assunto	2,148
7º	<i>term</i>	termo	1,627
8º	<i>system</i>	sistema	1,605
9º	<i>analysis</i>	análise	1,359
10º	<i>category</i>	categoria	1,090
11º	<i>concept</i>	conceito	1,086
12º	<i>user</i>	usuário	1,032
13º	<i>relationship</i>	relação	1,004
14º	<i>domain</i>	domínio	951
15º	<i>class</i>	classe	882
16º	<i>tag</i>	etiqueta; marcação	832
17º	<i>facet</i>	faceta	828
18º	<i>language</i>	linguagem	816
19º	<i>process</i>	processo	808
20º	<i>web</i>	rede (<i>internet</i>)	718

Fonte: Rocha (2023).

Esse contraste entre a Arquivologia e a Biblioteconomia resulta, de acordo com Cox *et al.* (2001) e Venancio (2017), da individualização da Arquivologia no cenário norte-americano em relação ao campo da Ciência da Informação. As principais razões apresentadas são: a autoafirmação como uma área do conhecimento, a ênfase do arquivista no documento em detrimento da informação e a existência posterior da Ciência da Informação em relação à Arquivologia (Venancio, 2017).

Essas características exemplificam o distanciamento dos arquivos nos interesses da comunidade discursiva em questão, dado que, apesar de a Arquivologia estar atualmente em um processo de aproximação da CI (no contexto norte-americano), há um resquício histórico de um afastamento nas últimas décadas. Ademais, outro aspecto constatado mais adiante no trabalho, é o direcionamento da comunidade para estudos de processos e ferramentas de organização da informação exclusivamente no âmbito das bibliotecas.

Subject (“assunto”, em português) ocorreu em 6ª posição. A partir de uma pesquisa por meio das ferramentas *Concordance* e *Word Sketch*, verificou-se que a palavra estabeleceu relações com “cabeçalhos”, “Cabeçalhos de Assunto da Biblioteca do Congresso”, “indexação”, “análise” e “representação”. Os termos *organization* (“organização”, 2,472) e *system* (“sistema”, 1,605) também devem ser destacados devido à relação direta dos termos com o objetivo da análise. Os demais termos evidenciaram aspectos diversos, como a *analysis* (“análise”, 1,359 ocorrências) que estabeleceu relações colocacionais, principalmente com os termos *facet* (148 colocações) e *domain* (118 colocações). Ou então, *tag* (832 ocorrências) e *web* (718 ocorrências) que demonstraram uma inclinação aos aspectos da organização do conhecimento/informação na *internet*.

Smiraglia (2007, p. 69, tradução nossa, grifo nosso) apresentou uma análise inicial da NASKO em seu artigo intitulado “A Glimpse at Knowledge Organization in North America” e constatou uma “[...] proeminência dos conceitos de **classificação facetada** e também um foco autoconsciente em todas as coisas novas, incluindo a **classificação social**”. Nesse ponto, é possível notar uma consistência que permanece ao longo dos anos. Sobretudo, na *classificação* que ocupou espaço entre os termos de maior frequência na *Wordlist*. Esse trecho do Smiraglia evoca

outra característica bastante marcante no decorrer da análise dessa comunidade discursiva, o contraste entre o tradicional e o “novo”.

5.2 Keywords

A ferramenta em questão, além de gerar dois tipos de listas (*Single-words* e *Multi-words terms*), possui uma abordagem voltada para a análise semântica. Utilizou-se de um *corpus* referencial (fornecido pelo próprio programa) para realizar o levantamento da frequência. Assim sendo, em vez de realizar uma análise métrica somente do *corpus* de estudo, a ferramenta faz uma comparação com outro conjunto textual utilizado como parâmetro. O *corpus* referencial selecionado foi o “English Web 2021 (enTenTen21)”, a versão mais recente entre os *corpora English Ten Ten*, com 52 bilhões de palavras extraídas de documentos na *web* (Sketch Engine, [202-?]a). Dentre os *corpora* disponíveis, esse foi o mais adequado devido a sua amplitude temática e variedade de estilos de escrita.

Entretanto, alguns critérios levantados por Sardinha (2004) para escolha do *corpus* de referência devem ser pontuados. Um *corpus* de referência que contenha uma grande variedade de gêneros não é o mais indicado para o estudo de *Keywords* (Sardinha, 2004). Isso se deve à consonância temática que deve existir entre os dois *corpora* para evitar resultados genéricos. Contudo, o autor também declara que o tamanho do *corpus* referencial influencia a quantidade dos resultados que podem ser obtidos. Sardinha (2004, p. 102) recomenda que o *corpus* de referência seja “[...] cinco vezes o tamanho do corpus de estudo”. Então o critério que norteia a escolha pelo “English Web 2021 (enTenTen21)” é, especialmente, a extensão desse *corpus*. Dessa forma, a limpeza de dados se concentrou em eliminar todas as generalidades que não se enquadraram no cerne temático.

Estabeleceu-se o lexema enquanto a classe de palavra principal de análise na *Keywords*, tanto para a *Single-words* como a *Multi-words terms*. A seguir, serão descritos e analisados os resultados submetidos ao mesmo processo de limpeza de dados aplicado à *Wordlist*.

5.2.1 Single-words

A listagem em questão contém palavras-chave constituídas de apenas uma palavra. Estão também inclusas as siglas, que excepcionalmente são palavras que carregam uma sentença. A análise se concentrou na comparação entre os dois *corpora* (“NASKO 2007-2021” e “English Web 2021 (enTenTen21)”). O valor em evidência é o *score* dos termos, considerando que esse é o total obtido da relação estabelecida entre ambos. Observe a tabela 2.

Tabela 2 - Lista de palavras-chave (*single-words*) do *corpus* da NASKO (2007-2021)

Ordem	Palavra	Tradução	Frequência absoluta		Score
			NASKO	Referência	
1º	<i>cataloge</i>	catalogar	687	83,511	389.7
2º	<i>ddc</i>	CDD	391	24,746	373.1
3º	<i>lcsb</i>	Cabeçalhos de assunto da Biblioteca do Congresso	274	3,357	347.9
4º	<i>koss</i>	SOCs	263	7,601	313.4
5º	<i>bibliographic</i>	bibliográfico	547	83,094	311.3
6º	<i>indexing</i>	indexação	688	144,957	274.2
7º	<i>frbr</i>	Requisitos Funcionais para Registros Bibliográficos	211	4,444	263.7
8º	<i>classification</i>	classificação	3,398	1,150,480	230.6
9º	<i>ko</i>	OC	692	214,882	206.0
10º	<i>classificatory</i>	classificatório	149	3,588	188.9
11º	<i>ontology</i>	ontologia	543	206,819	166.6
12º	<i>bibliographical</i>	bibliográfico	189	33,170	164.7
13º	<i>cataloger</i>	catalogador	145	12,450	161.9
14º	<i>kos</i>	SOC	292	93,566	155.1
15º	<i>facet</i>	faceta	828	379,290	154.6
16º	<i>thesaurus</i>	tesauro	215	57,731	148.7
17º	<i>taxonomy</i>	taxonomia	354	169,603	126.2
18º	<i>indexer</i>	indexador	126	22,052	124.6
19º	<i>headings</i>	cabeçalhos	116	16,626	122.7
20º	<i>rda</i>	RDA	152	45,520	117.3

Fonte: Rocha (2023).

Cataloge (“catalogar”, 389.7) foi a primeira palavra que surgiu na listagem e na décima segunda posição, *cataloger* (“catalogador”, 161.9). Alguns conceitos presentes nos processos de representação descritiva e temática, tais como CDD (2ª posição, 373.1), LCSH (3ª posição, 347.9), FRBR (6ª posição, 263.7) e o RDA (20ª posição, 117.3), também apareceram na listagem. Tradicionalmente, a prática da catalogação tem sido influenciada pela tradição anglo-americana moderna que surgiu de bibliotecas individuais no Reino Unido e Estados Unidos (EUA), promovendo uma padronização internacional em países falantes da língua inglesa (Dobreski, 2020). Tanto para Dobreski (2020) como para Cristina Aganette, Teixeira e Pinto Aganette (2017), os padrões de catalogação descritiva possuem uma presença significativa no desenvolvimento da atividade de catalogação.

O RDA é um código de catalogação desenvolvido a partir de modelos conceituais, entre eles o FRBR, que surgiu como uma proposta de substituição ao Código de Catalogação Anglo-Americano (AACR2) (Cristina Aganette; Teixeira; Pinto Aganette, 2017). Panchyshyn, Lambert e McCutcheon (2019, p.120, tradução nossa) afirmam que em 2017, o “[...] RDA era implementado pela maioria das principais bibliotecas acadêmicas e de pesquisa dos EUA”, entre elas, as três bibliotecas nacionais: “[...] Library of Congress (LC), a National Library of Medicine (NLM), e a National Agricultural Library (NAL)”. Portanto, os dados apresentam uma inclinação aos estudos de normas, códigos e processos atrelados a catalogação que são característicos da cultura das bibliotecas norte-americanas.

Ainda, a respeito da representação centrada no conteúdo, as atividades de *indexing* (“indexação”, 274.2) e *classification* (“classificação”, 230.6) estiveram, respectivamente, em 6º e 8º lugar. A profissão do indexador (*indexer*, 124.6) apareceu entre os últimos na décima oitava posição. Com isso, conclui-se que, para os integrantes da ISKO-C/US, as práticas de representação descritiva e temática são díspares entre si. O catalogador descreve o documento, enquanto o indexador representa o seu conteúdo. Além disso, foi visível que a ênfase semântica do *corpus* da NASKO está mais centrada na descrição do que na temática. A diferenciação entre esses dois tipos de representação pode ser descrita pela perspectiva de Hjørland (2008) como práticas de organização de conhecimento *sintáticas* e *semânticas*. Em outras palavras, a representação descritiva está para sintaxe, assim

como a temática está para a semântica. Com base nessa análise, a comunidade tem se preocupado mais com as atividades sintáticas da área.

De acordo com Guimarães (2008, p. 106), sob a perspectiva do universo do Tratamento Temático da Informação (TTI)¹⁰, a tendência da comunidade discursiva da NASKO para a catalogação é influenciada pela vertente da catalogação de assunto (origem norte-americana). A indexação (inglesa) e a análise documental (francesa) têm um impacto menor no *corpus*.

No que diz respeito aos Sistemas de Organização do Conhecimento, a sua sigla se posicionou na quarta posição (“koss”, 313.4) e dentre os SOC de maior interesse da comunidade estiveram: *ontology* (“ontologia”, 166.6) na 10ª posição, *thesaurus* (“tesauro”, 148.7) na 16ª posição e *taxonomy* (“taxonomia”, 126.2) na 17ª posição. Então, dentre os objetos de interesse, a ontologia foi considerado um dos mais relevantes para essa comunidade discursiva. Como mencionado anteriormente, dentre os SOC analisados, a ontologia é a que apresenta maior riqueza semântica entre eles. Hjørland (2008, p. 120, tradução nossa) explica que as ontologias “[...] são tipos de relações semânticas ilimitadas” e “[...] produzidas para fazer inferências lógicas por computadores e coloca, portanto, altas exigências nas especificações formais”.

De 2007 até 2021, foi constatada uma progressão das ocorrências do termo *ontology* nos anais da NASKO. Inicialmente, em 2007 e 2009, a frequência era de cinco publicações mencionando o tópico. A partir de 2011, houve um aumento de 10 vezes que o termo aparece nos artigos. No último ano de 2021, o termo foi mencionado em 11 dessas publicações. É importante ressaltar que a média de publicações transita entre 10 a 20 artigos/resumos expandidos a cada ano de evento. Essa progressão indica que a comunidade discursiva está em constante atualização em relação a esse tópico. Demonstrando que a NASKO acompanhou o desenvolvimento das discussões acerca das ontologias ao longo desse período.

5.2.2 Multi-words terms

A tabela 3 apresenta a listagem de palavras-chave constituídas de mais de uma palavra. Essa característica, por si só, é responsável por apresentar resultados

¹⁰ Cf. Guimarães (2008, p. 105-106).

distintos em comparação com a lista anterior. Isso ocorre, pois, o *software* opera identificando com precisão o que é um termo isolado e aqueles que são compostos.

O primeiro termo evidenciou a temática central da comunidade, a *knowledge organization* (“organização do conhecimento”, 2,094.0). Observe o tabela 3.

Tabela 3 - Lista de palavras-chave (Multi-words terms) do *corpus* da NASKO (2007-2021)

Ordem	Palavra	Tradução	Frequência absoluta		Score
			NASKO	Referência	
1º	<i>knowledge organization</i>	organização do conhecimento	1,681	4,449	2,094.0
2º	<i>subject headings</i>	cabeçalhos de assunto	277	15,964	294.5
3º	<i>information science</i>	ciência da informação	401	65,723	259.5
4º	<i>controlled vocabulary</i>	vocabulário controlado	185	10,882	210.8
5º	<i>facet analysis</i>	análise de facetas	154	324	205.5
6º	<i>classification scheme</i>	esquema de classificação	172	14,760	186.0
7º	<i>decimal classification</i>	classificação decimal	139	1,480	182.2
8º	<i>knowledge organization system</i>	sistema de organização do conhecimento	133	994	175.7
9º	<i>domain analysis</i>	análise de domínio	135	2,745	173.5
10º	<i>subject indexing</i>	indexação de assunto	126	966	166.6
11º	<i>library classification</i>	classificação de biblioteca	111	918	147.0
12º	<i>bibliographic classification</i>	classificação bibliográfica	107	218	143.3
13º	<i>information retrieval</i>	recuperação de informação	170	38,894	139.7
14º	<i>classification system</i>	sistema de classificação	165	41,398	132.3
15º	<i>indexing language</i>	linguagem de indexação	98	421	130.9
16º	<i>faceted classification</i>	classificação facetada	95	886	126.0
17º	<i>subject analysis</i>	análise de assunto	89	738	118.4
18º	<i>classification theory</i>	teoria da classificação	84	403	112.4
19º	<i>dewey decimal</i>	decimal de dewey	86	3,146	110.2
20º	<i>metadata schema</i>	esquema de metadados	83	3,059	106.5

Fonte: Rocha (2023).

No segundo lugar, foi notada a presença do sistema de organização do conhecimento, *subject headings* (“cabeçalhos de assuntos”, 294.5). De acordo com Guimarães (2008, p. 106), os cabeçalhos de assunto e a catalogação alfabética de Cutter; são, tradicionalmente, pertencentes à vertente teórica da catalogação de assunto e têm sua origem na *Library of Congress*, “[...] cuja ênfase reside no catálogo enquanto produto do tratamento da informação em bibliotecas (subject cataloguing)”. Em terceiro lugar, observou-se a *information science* (“ciência da informação”, 259.5). Os principais Sistemas de Organização do Conhecimento (na oitava posição, *knowledge organization system*, 175.7) presentes nessa conjuntura foram, *subject headings* e *classification scheme* (“esquema de classificação”, 186.0).

Controlled vocabulary (“vocabulário controlado”, 210.8) e *indexing language* (“linguagem de indexação”, 130.9), posicionaram-se, respectivamente, em quarto e décimo quinto lugar. Conforme Rubi (2009), a indexação é considerada uma atividade subjetiva e, portanto, para identificar e representar um assunto necessita de um controle terminológico. Com o propósito de evitar uma dispersão da recuperação do objeto documental registrado (que inclusive está presente na 13ª posição através do *information retrieval*, “recuperação da informação” com 139.7 de score). Tendo em vista isso, percebeu-se que no escopo da indexação, o vocabulário controlado é o tema mais usual. Além disso, outras atividades relacionadas devem ser mencionadas, tais como: *subject indexing* (“indexação de assunto”, 166.6) e *subject analysis* (“análise de assunto”, 118.4). Com isso, percebeu-se a presença da vertente de origem inglesa da indexação no *corpus*.

Michael Buckland, em 1996, traz uma explicação para o desinteresse na abordagem da análise documentária em seu artigo “Documentation, Information Science, and Library Science in the U.S.A”, ao apontar a relutância dos Estados Unidos ao documentalismo europeu após a Segunda Guerra Mundial. Buckland (1996) expôs que havia um interesse de desenvolvimento tecnológico entre a Europa e EUA, na virada do século XIX até a Segunda Guerra Mundial, porém não havia uma partilha de ideias recíproca entre os dois grupos. Na Ciência da Informação, em especial — apesar das contribuições significativas — os documentalistas europeus foram ignorados pelos estadunidenses. Esse traço histórico mencionado serve, por exemplo, para compreender dentro do *corpus* a pouca relevância do enunciado-pivô “Documentary Representation” abordado.

Deve-se ressaltar que nessa lista de 20 termos, *classification* (“classificação”) apareceu sete vezes nas seguintes posições: 6^a ***classification scheme***, 7^a ***decimal classification*** (“classificação decimal”, 182.2), 11^a ***library classification*** (“classificação de biblioteca”, 147.0), 12^a ***bibliographic classification*** (“classificação bibliográfica”, 143.3), 14^a ***classification system*** (“esquema de classificação”, 132.3), 16^a ***faceted classification*** (“classificação facetada”, 126.0), e 18^a ***classification theory*** (“teoria da classificação”, 112.4). Além disso, na décima nona posição apareceu *dewey decimal* que se refere à Classificação Decimal de Dewey. Com isso, chegou-se à conclusão de que a classificação é analisada nas suas mais diversas facetas. Desde a teoria geral até os tipos específicos, todos eles voltados, na maioria das vezes, para o contexto da biblioteca.

Reitera-se que tanto na etapa inicial de verificação dos enunciados-pivô quanto nos dados adquiridos por meio do *Sketch Engine*, “Representação Documentária” apresentou uma ocorrência rara dentro do *corpus*. Até esse ponto do trabalho, é possível identificar que a vertente norte-americana impera no conteúdo das publicações. Em seguida, está a vertente inglesa que reside na temática da indexação de assunto e classificação. Sendo a vertente francesa, isto é, a análise documental quase inexistente nas discussões.

5.3 Word Sketch

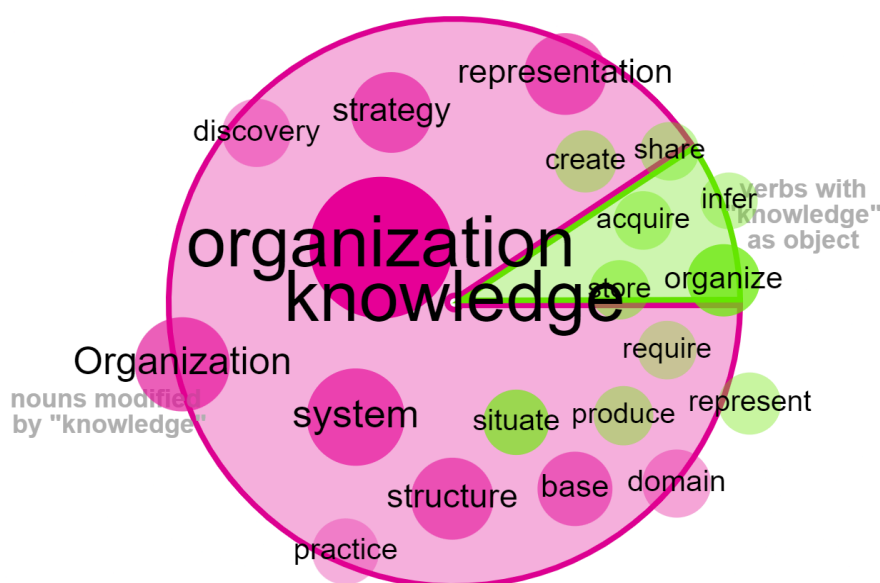
No *Word Sketch*, os gráficos abordam o *lemma* como termo central em relação às classes de palavras (colocações) *verbs* e *nouns modifiers*. Ao contrário dos quadros anteriores, os gráficos do *Word Sketch* e *Word Sketch Difference* são gerados pelo próprio *software*.

Não foi estabelecido recorte de análise para essa ferramenta, pois, de acordo com Sinclair (1991), as colocações recorrentes tendem a ter uma frequência muito baixa a depender da extensão textual. No entanto, o autor não descarta a possibilidade de existirem situações especiais que podem fornecer informações relevantes para análise (Sinclair, 1991). O que realmente ocorre com o *corpus* da NASKO, portanto, serão consideradas três situações: a distribuição das colocações no total de textos reunidos, as colocações que se sobressaem entre as demais e os

enunciados-pivô predefinidos. Os termos que aparecem nos gráficos e não são mencionados na análise, não contemplam as situações anteriormente elencadas.

O termo *knowledge* (“conhecimento”) apresentou uma relação significativa com *organization* (“organização”, 522 colocações). Considerando que a comunidade discursiva em questão se posiciona em um campo de especialidade científica — a organização do conhecimento —, a colocação corresponde com a representatividade temática central do *corpus*. Tal como apresentado na figura 1.

Figura 1 - *Noun modifiers e verbs with object (Knowledge)*



visualization by  SKETCH ENGINE

Fonte: Sketch Engine (2023).

Outra questão a ser analisada é o contraste verificado entre *organization* e *representation* (“representação”, 66 colocações). Foi constatada uma proeminência de *knowledge organization* em relação à *knowledge representation*, sendo a primeira (aproximadamente) oito vezes mais recorrente que a segunda. O mesmo se aplica aos verbos *organize* (“organizar”, 34 colocações) e *represent* (“representar”, 11 colocações).

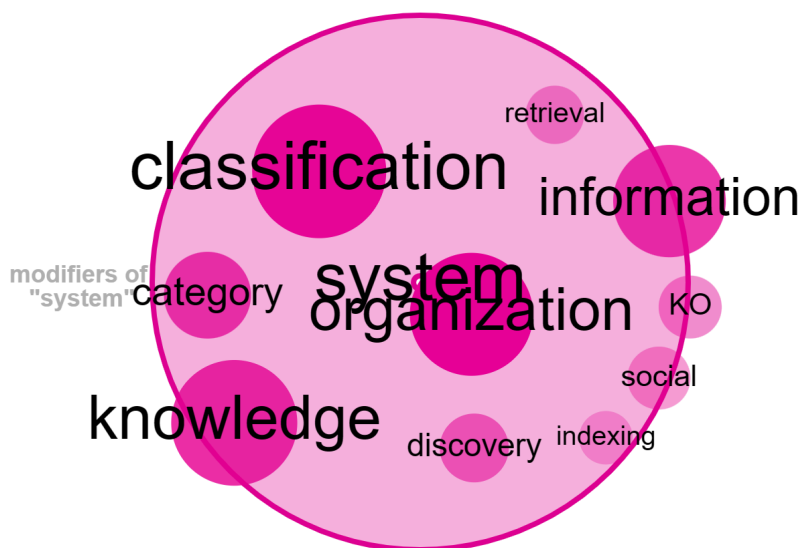
A comunidade discursiva utiliza como cerne temático a expressão “organização do conhecimento” para se referir a sua própria área de atuação. Para Hjørland (2008), a OC, em seu sentido específico, diz respeito a atividades como

indexação, classificação e descrição de documentos. Atividades essas, mencionadas constantemente nas publicações da NASKO. Por outro lado, Brascher e Café (2008), contestam essa definição, ao declarar que se refere à “organização da informação” e não à “organização do conhecimento”. Percebe-se então que essa distinção entre “organizar” e “representar” a informação ou o conhecimento é própria do contexto brasileiro.

A constatação foi corroborada pelo fato de que a maioria dos artigos com *knowledge representation* apresenta autores brasileiros ou fazem referência ao uso do termo no contexto brasileiro. Exceto pelas situações em que o termo é vinculado ao contexto tecnológico. Dado que a “representação do conhecimento”, de acordo com Jian Qin (2002), é proveniente da Ciência da Computação, relacionando-se especificamente com a Inteligência Artificial (IA). De acordo com Qin (2020), a organização do conhecimento pode ser dividida em duas áreas: o conhecimento organizado pela compreensão humana em sistemas (como, por exemplo, esquemas de classificação e tesouros) e a aplicação dos SOC (por intermédio da ação humana ou máquina) para representação do conteúdo de um documento. Portanto, para essa comunidade, a representação não é compreendida como uma oposição à organização, mas sim um elemento essencial no processo de aplicação dos Sistemas de Organização do Conhecimento.

A respeito da palavra *system* (“sistema”), foram analisados os *noun modifiers*. A aplicação da análise somente por uma classe lexical ocorreu para destacar quatro colocações relevantes, uma vez que não foram observadas frequências significativas com relação aos verbos presentes no termo analisado da figura 2.

Figura 2 - Noun modifiers (system)



visualization by  SKETCH ENGINE

Fonte: Sketch Engine (2023).

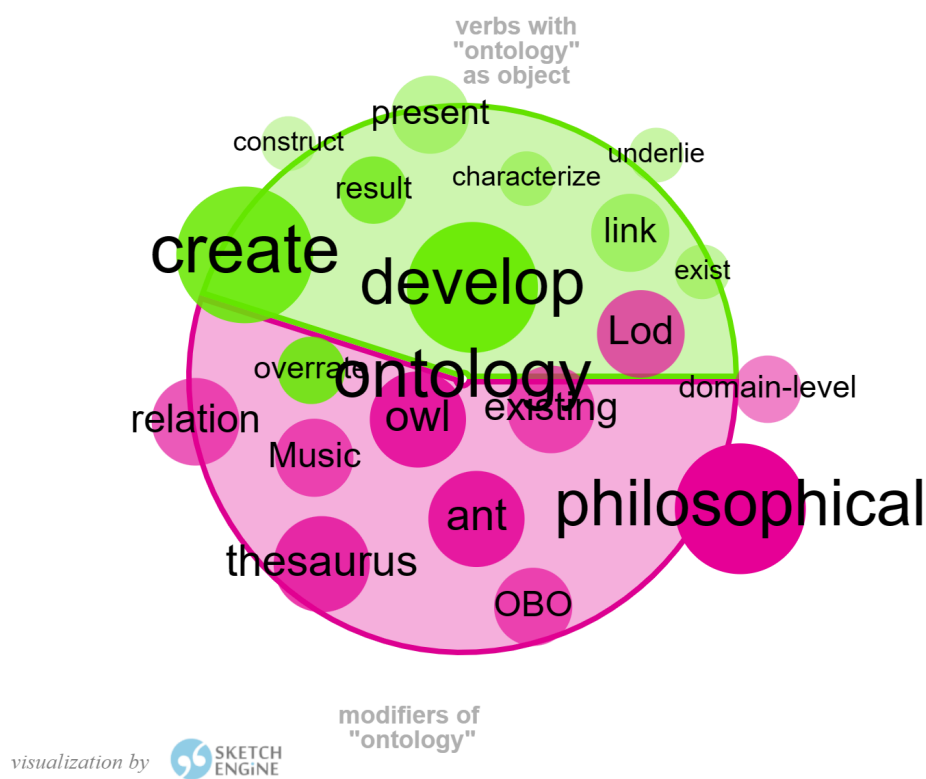
Dentre as quatro palavras, *organization* (“organização”) foi a que mais se destacou, com 137 colocações, sendo *[knowledge] organization systems* (123 colocações) o mais recorrente dessa relação. *Knowledge* (143 colocações), que está em terceiro lugar, também apresenta sua incidência relacionada ao termo *knowledge [organization] system*. A segunda palavra mais próxima de *system* é *classification* (161 colocações), retomando o interesse acerca da classificação já observada nas *Keywords*.

Por último, em quarto lugar, ocorreu *information* com 115 colocações. Uma análise individual indicou que as relações estabelecidas foram, consecutivamente: *information system* e *[knowledge] organization system*. A respeito de *information*, dentre as 115 colocações, identificaram-se as seguintes relações: *information [representation] system*, *information [discovery] system*, *information [retrieval] system*, *information [organization] system*. Os trechos em que “information representation system” recebeu menção se referem aos sistemas de organização do conhecimento, sobretudo as ontologias. *Information system* não é necessariamente um indicativo de uma sinonímia para SOC, e sim, visto como uma associação de implementação mútua entre os dois.

Conforme Hjørland (2015), o termo “Knowledge Organization System” tornou-se comum em 2000, através do relatório de Gail Hodge. Com base nisso, analisou-se que o enunciado-pivô foi mencionado em apenas três artigos de um total de 14 publicados em 2007 nos anais da NASKO. O que demonstrou que naquele ponto inicial, a temática ainda não era tão popular entre os pesquisadores da comunidade. Houve um aumento da utilização do termo, a partir do segundo evento ocorrido em 2009 e nos anos seguintes até 2021.

Os SOC *ontology* (“ontologia”), *thesaurus* (“tesauro”) e *taxonomy* (“taxonomia”), foram analisados individualmente com suas respectivas colocações. Começando pela ontologia da figura 3.

Figura 3 - *Noun modifiers* e *verbs with object* (*ontology*)



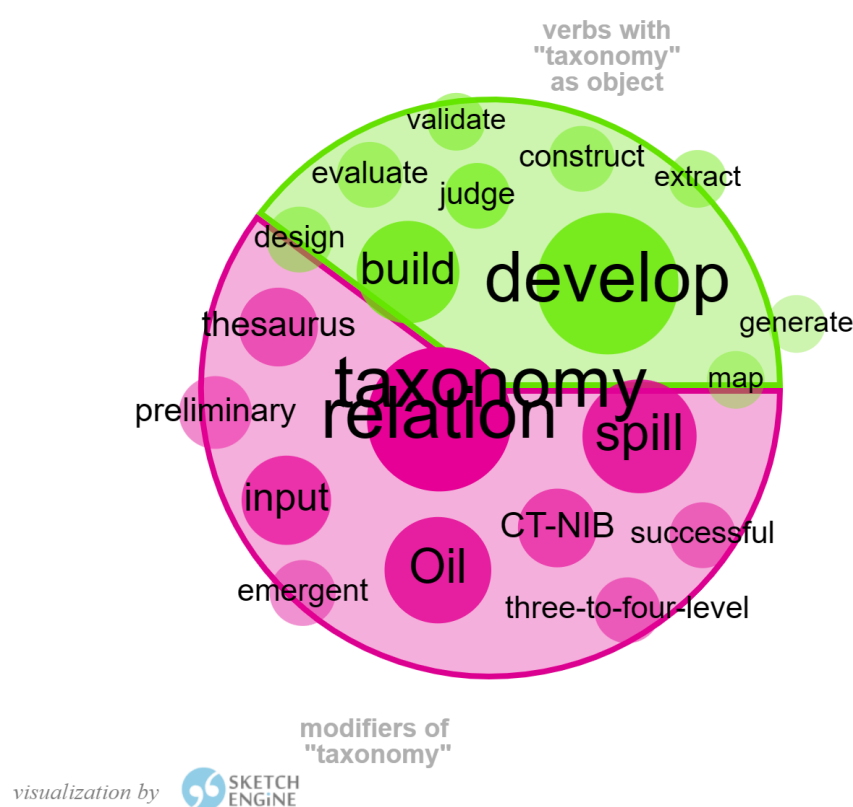
Fonte: Sketch Engine (2023).

Verificou-se que o verbo é a classe gramatical que frequentemente estabeleceu colocações com ontologia, ao contrário dos resultados apresentados até esse ponto com os demais termos. Foram os casos de maior frequência: *develop* (“desenvolver”, 11 colocações) e *create* (“criar”, 12 colocações). O que indicou uma tendência maior do termo para ações e atividades do que qualificações. No que diz

respeito aos *nouns modifiers*, a principal relação foi estabelecida com *philosophical* (“filosófico”, 11 colocações). As demais colocações não apresentam números significativos, entretanto apresentam estabelecimento com teorias e linguagens ontológicas como, **ANT ontology** (“Teoria Ator-Rede”), **OWL ontology** (“Linguagem de Ontologia Web”), **LOD ontology** (“Linked Open Data”) e **OBO ontology** (“Open Biological and Biomedical Ontologies”).

A palavra *taxonomy* não estabeleceu distribuições significativas ao decorrer do *corpus*. Como, por exemplo, o verbo *develop* (22 colocações distribuídas em um artigo) e o *noun modifier relation* (“relação”, 23 colocações distribuídas em dois artigos). Observe a figura 4.

Figura 4 - *Noun modifiers* e *verbs with object (taxonomy)*



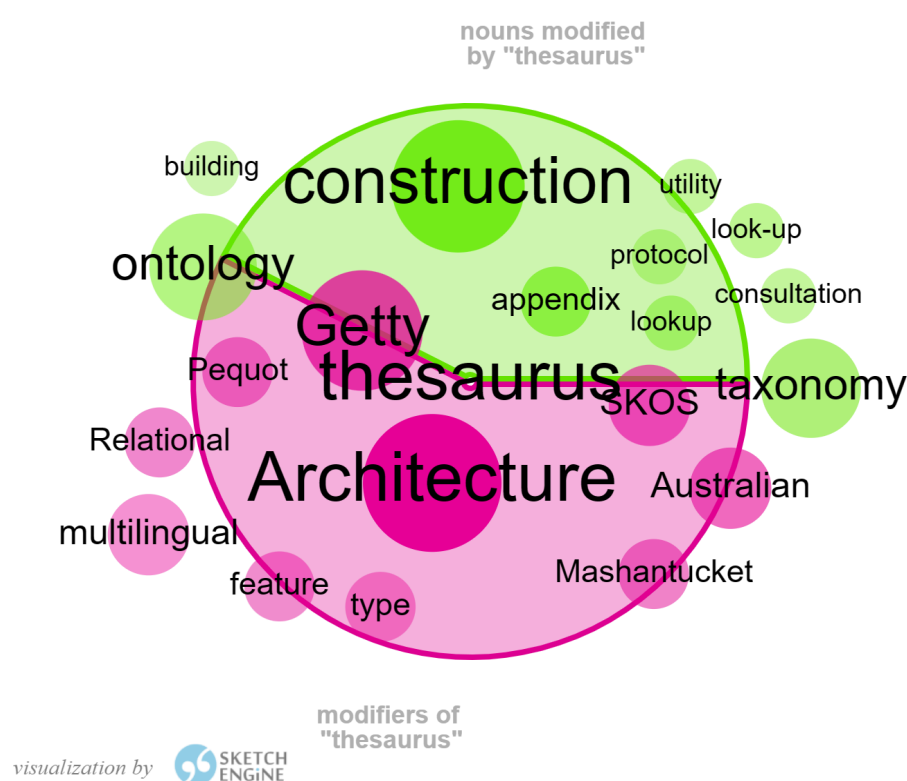
Fonte: Sketch Engine (2023).

Observou-se que as colocações frequentes ocorreram pontualmente em dois ou um dos artigos do *corpus*. Dessa forma, não foram detalhadamente analisadas, contudo, mesmo as escassas relações estabelecidas são um comportamento

indicativo. Dado que demonstra uma abordagem desse termo menos frequente nos tópicos dessa comunidade.

O termo *thesaurus* estabeleceu maior ocorrência com o *noun modifier construction* (“construção”, 10 colocações). Apesar do número baixo de colocações, houve uma distribuição uniforme dessa colocação por nove artigos. Outra característica observada, apesar da quantidade baixa de ocorrências, é a predominância de nomes próprios de tesouros na figura 5.

Figura 5 - *Noun modifiers e verbs with object (thesaurus)*



Fonte: Sketch Engine (2023).

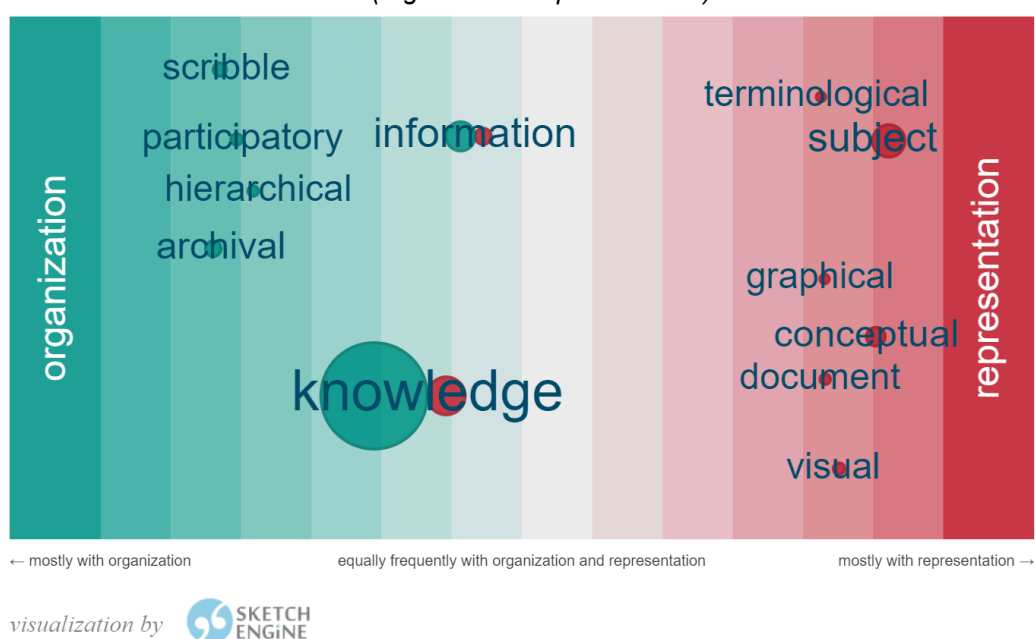
Os principais tesouros encontrados foram, *Art & Architecture Thesaurus* (AAT), *Getty Thesaurus of Geographic Names*, *Australian Thesaurus of Education Descriptors* (ATED) e *The Mashantucket Pequot thesaurus of American Indian terminology project*. As relações em si não são recorrentes, contudo, ao reunir essas colocações, percebe-se novamente uma tendência da comunidade discursiva em debruçar-se em estudos de modelos/projetos específicos.

5.4 Word Sketch Difference

A ferramenta em questão também analisa colocações, mas o faz de forma comparativa entre dois lexemas. Primeiro, são pesquisadas duas palavras e, a partir disso, um gráfico é gerado com as colocações específicas de cada enunciado-pivô e suas colocações em comum. Enfim, aplica-se um filtro que dispõe de um conjunto de classes de palavras para estabelecimento das relações entre os dois enunciados-pivô selecionados e suas respectivas colocações. Sendo necessária uma apropriação da terminologia da área e da língua para estabelecer esses critérios lexicais e/ou gramaticais.

O primeiro gráfico utilizou o filtro “modifiers of...”, para análise dos qualificadores dos termos *organization* e *representation*. A figura 6 evidencia que *knowledge* e *information* obtiveram maior inclinação colocacional para a organização quando comparado com a representação.

Figura 6 - *Word Sketch Difference (organization/representation)*



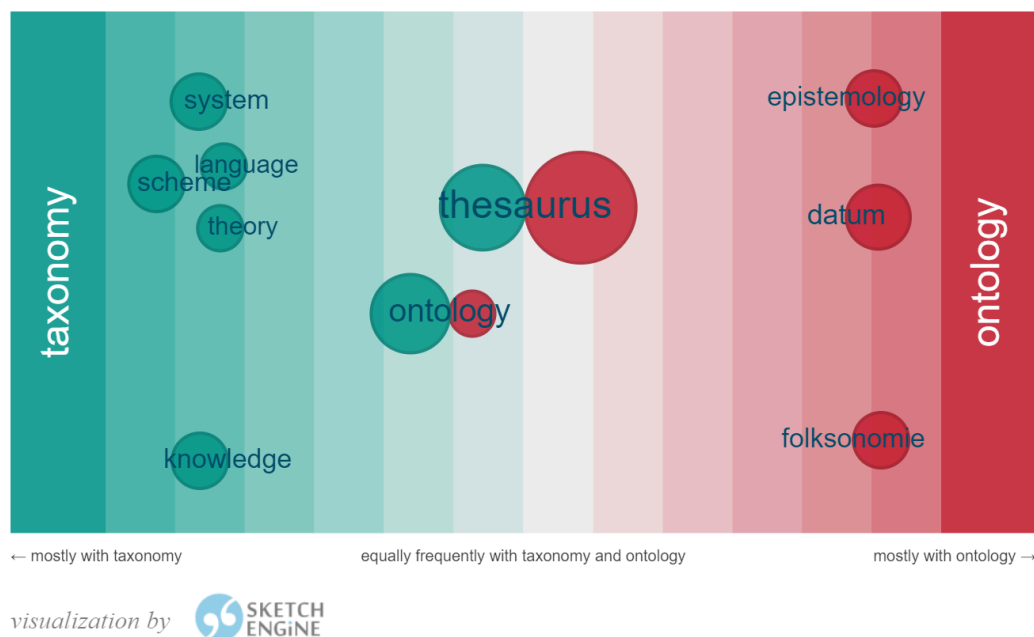
Fonte: Sketch Engine (2023).

Salienta-se, contudo, que *representation* estabeleceu maior ocorrência com *subject* (“assunto”, 48 colocações) e *conceptual* (“conceitual”, 16 colocações). Essas relações demonstraram a presença implícita da representação do conhecimento,

matéria que considera aspectos como classificação, conceitos, relacionamentos e características (Brascher; Café, 2008, p. 7). O que evidencia novamente que a representação e organização do conhecimento almeja, por meio da instituição dos seus estudos e atividades, a recuperação do conhecimento. Hjørland (2009) explica que os conceitos estão por trás das questões informacionais dos usuários e que se relaciona diretamente com a organização do conhecimento para desenvolvimento de Sistemas de Organização do Conhecimento.

Outro aspecto marcante observado durante a primeira análise com os enunciados-pivô é a intercalação entre os termos *thesaurus*, *taxonomy* e *ontology*. Por isso, no gráfico de comparação entre *taxonomy* e *ontology*, foi aplicado o filtro “and/or...” (“e/ou”, em português). Não há uma fórmula de análise dos termos nas relações colocacionais, visto que, sob a perspectiva empírica, sabe-se de relações mais ou menos proeminentes. Principalmente no que diz respeito à terminologia de uma área científica. Esse aspecto deve ser destacado para compreender o motivo pelo qual as análises colocacionais têm abordagens de classe lexicais distintas. Durante este processo, observou-se que os SOC foram mencionados de maneira listada, tal como pode ser observado na figura 7.

Figura 7 - *Word Sketch Difference (taxonomy/ontology)*



Fonte: Sketch Engine (2023).

Dessa forma, comprovou-se que tesouros, ontologias e taxonomias têm uma relação concomitante. Não foi identificada nenhuma relação comparativa entre os termos, e sim uma relação igualitária. Com base em estudos de autores como Zeng (2008), Hjørland (2015) e Barros, Bastos e Santos (2022), é de conhecimento geral que Sistemas de Organização do Conhecimento possuem características determinantes para aplicações específicas. Logo, os estudos da comunidade discursiva acerca da temática não se baseiam nas vantagens em detrimento de outras, mas sim no contexto em que cada uma se aplica.

Ainda é importante lembrar que além desses três SOC foram identificados ao longo da análise mais outros abordados pela comunidade, os *esquemas de classificação* e *cabeçalhos de assuntos*. O que corrobora para a uniformidade do trabalho realizado, dado que esses SOC apresentaram uma recorrência semântica. Além disso, retoma as principais abordagens do tratamento temático apresentadas no *corpus* até então, as vertentes norte-americana e inglesa.

6 CONSIDERAÇÕES FINAIS

Após uma análise mais aprofundada desses dados, foi perceptível a existência de um contraste entre o avanço teórico de novos métodos e abordagens, ao mesmo tempo que a comunidade discursiva estabelece uma credibilidade ao que é tradicionalmente fundamentado na área da OC. No entanto, até mesmo o que é considerado tradicional pode ser alvo de ressalvas e aperfeiçoamentos. Observou-se, por exemplo, que a CDD é alvo de uma análise crítica, uma vez que apresenta problemas pendentes de solução, de acordo com alguns membros da ISKO-C/US. Problemas tais como, a necessidade de uma classificação mais ampla para outros grupos étnicos, revisão do sistema para inclusão dos povos originários norte-americanos, expansão da classe 200 (Religião) para promoção da diversidade religiosa no sistema e análise do papel de gênero da mulher na CDD. Essas demandas vão de encontro com a análise realizada por Hjørland (2008), que pontua que a CDD foi projetada para fins administrativos do bibliotecário e não para o usuário.

Outro fator relevante foi a consistência do uso dos conceitos em relação aos temas abordados pela comunidade. As variações, quando existem, desempenharam pouca influência no perfil terminológico (*Wordlist* e *Keywords*) do conjunto textual da comunidade discursiva da NASKO. Um exemplo a ser destacado é a apropriação sistemática do termo *Knowledge Organization* para se referir à sua própria disciplina. Esse padrão possibilita concentrar o progresso teórico no campo em questões urgentes, ao mesmo tempo, em que pode limitar as discussões dos conceitos que poderiam promover favorecimento ao desenvolvimento científico. A partir dessa perspectiva, é possível refletir em como essas decisões reverberam no *corpus*. Com base nisso, é possível inferir que há um perfil terminológico que pode ser seguido, mas não restringe os estudos de pesquisadores de outros países (como foi observado com os pesquisadores brasileiros). Conforme mencionado, diversas vezes ao longo deste trabalho, a comunidade discursiva se preocupa com as questões atuais que cercam a organização do conhecimento e estão aperfeiçoando constantemente as suas discussões.

Ademais, concluiu-se que o analista deve possuir domínio do idioma de uma comunidade discursiva posicionada em outro contexto social e histórico. O que

requer a tradução das principais características de um dado plano cultural para outro de forma compreensível. A exemplo disso, estão os termos “representação descritiva” e “representação temática” que são próprios da OC no contexto brasileiro e, portanto, são abordadas através de termos equivalentes propostos por Hjørland (2008) e Guimarães (2008) nas análises das *Keywords (single-words)*. Nesse caso, o analista precisa se apropriar de duas línguas e dois contextos para análise dos enunciados. Ou seja, aplicar a análise de discurso e se opor a um modelo puramente lógico de descrição e interpretação (Pêcheux, 1997).

Houve uma predominância de atividades que se concentram na organização de itens bibliográficos, próprios de uma biblioteca. O que ocorre devido a significativa influência da comunidade da organização do conhecimento norte-americana em desenvolver códigos, normas e ferramentas para a organização bibliográfica. Além disso, há o distanciamento entre a Documentação europeia e o campo da Ciência da Informação norte-americana, e também foram observados aspectos influenciados pela indexação, que é uma vertente inglesa. O que levanta outro aspecto a ser refletido, a importância da incorporação da língua inglesa nas práticas de organização do conhecimento para autonomia dos profissionais e pesquisadores na sua própria área de atuação. Se, por exemplo, o catalogador ou o indexador não se apropriarem das matérias na sua origem, estão sujeitos aos enganos causados pelas traduções ou pela dificuldade de recuperação do documento original. Essa é uma questão a ser pensada, considerando o grau de influência e adoção que essas ferramentas e teorias têm na organização do conhecimento internacional.

No que diz respeito à representação, alguns elementos se destacaram. Entre esses o contraste entre a “organização” e a “representação” do conhecimento, que era previsto na análise. Todavia, observou-se em seu sentido explícito uma influência relacionada à Ciência da Computação, oriunda dos pesquisadores norte-americanos da NASCO. Assim como as ontologias, que foram citadas diversas vezes em conjunto com temáticas relacionadas às tecnologias. A partir da preposição *of*, o enunciado “representação” estabelece relações significativas com a colocação “assunto”. No sentido implícito da representação predomina a menção por ferramentas (código de catalogação e modelo conceitual) de cunho descritivo.

A ontologia apareceu como Sistema de Organização do Conhecimento mais recorrente entre os enunciados-pivô pré-estabelecidos para a análise. Além disso, este SOC apresentou uma maior distribuição de suas colocações pelo *corpus* em comparação com os outros sistemas (taxonomia e tesouro). Embora o aspecto filosófico da ontologia tenha maior recorrência, houve uma relação com atividades de criação e desenvolvimento, no que diz respeito às colocações verbais. Além das taxonomias e tesouros, também foram encontrados os SOC “cabeçalhos de assunto” e “esquemas de classificação”, que estão relacionados, respectivamente, aos Cabeçalhos de Assunto da Biblioteca do Congresso e a Classificação Decimal de Dewey. De modo geral, o interesse identificado em avaliar e teorizar acerca dos SOC indica, mais uma vez, o caráter do autoaperfeiçoamento e atualidade presentes nas discussões teóricas da NASKO.

Durante a análise das *Keywords*, é importante salientar que foi constatada a necessidade de um *corpus* referencial na temática da Organização do Conhecimento em inglês, assim como é orientado por Sardinha (2004). Esse pode ser um ponto a ser considerado mais adiante, uma vez que esta análise integra um projeto de pesquisa em andamento. Assim, espera-se agregar um maior aprofundamento através da perspectiva norte-americana. Não há pretensão de promover uma terminologia unívoca, mas de organizar as conceituações a fim de identificar lacunas, comparar pontos de vistas e aperfeiçoar discussões na área através das produções científicas internacionais.

A partir de estudos como esse, será possível vislumbrar desdobramentos e assuntos que serão mais pontuais para as comunidades que integram a ISKO. A NASKO, por exemplo, tem inclinação para se dedicar aos Sistemas de Organização do Conhecimento (com destaque para as ontologias) e a formular revisões/proposições de ferramentas para a organização da informação. Espera-se também fornecer um aporte para emprego de teorias metodológicas centradas na representação e a definir, de maneira fundamentada, as perspectivas conceituais da organização no contexto norte-americano.

REFERÊNCIAS

ALMEIDA, Josiana Florêncio Vieira Régis de; DIAS, Guilherme Ataíde. Estado da arte sobre análise de domínio no campo da Ciência da Informação brasileira. **Brazilian Journal of Information Science: Research Trends**, Marília, v. 13, n. 2, p. 26-45, 2019. DOI: 10.36311/1981-1640.2019.v13n3.04.p26. Disponível em: <https://revistas.marilia.unesp.br/index.php/bjis/article/view/8610>. Acesso em: 31 out. 2023.

AZEREDO, José Carlos Santos de. **Iniciação à sintaxe do português**. Rio de Janeiro: Zahar, 2012. *E-book*. Disponível em: beta.lectio.com.br. Acesso em: 24 jul. 2023.

BARITÉ, Mario. **Diccionario de organización del conocimiento**: clasificación, indización, terminología. Montevideo: CSIC, 2015. *E-book*. Disponível em: <https://hdl.handle.net/20.500.12008/9028>. Acesso em: 2 ago. 2023.

BARROS, Thiago Henrique Bragato. A análise de discurso de Michel Pêcheux e a organização do conhecimento: possibilidades teórico-metodológicas. *In*: AMORIM, Igor Soares; SALES, Rodrigo de (org.). **Ensaio em organização do conhecimento**. Florianópolis: UDESC, 2021. p. 61-78. Disponível em: <https://dx.doi.org/10.5965/978-65-88565-20-9>. Acesso em: 19 jul. 2023.

BARROS, Thiago Henrique Bragato. Michel Pêcheux's Discourse Analysis: an approach to Domain Analyses. **Knowledge Organization**, [s. l.], v. 50, n. 1, p. 3-10, 2023. DOI: 10.5771/0943-7444-2023-1-3. Disponível em: <https://www.nomos-elibrary.de/10.5771/0943-7444-2023-1/ko-knowledge-organization-jahrgang-50-2023-heft-1?page=1>. Acesso em: 9 jan. 2024.

BARROS, Thiago Henrique Bragato; BASTOS, Carine Melo Cogo; SANTOS, Adriana Carla Ribeiro dos. Sistemas de Organização do Conhecimento no contexto da Arquivologia: aportes metodológicos para seu desenvolvimento. **Revista Acervo**, [s. l.], n. 2, v. 35, p. 1-20, maio/ago. 2022. Disponível em: <https://revista.an.gov.br/index.php/revistaacervo/article/view/1812/1726>. Acesso em: 31 jul. 2023.

BARROS, Thiago Henrique Bragato; LAIPELT, Rita do Carmo Ferreira. Uma análise de domínio da área de Organização e Representação do Conhecimento no contexto do periódico *Em Questão*. **Em questão**, Porto Alegre, v. 27, n. 4, p. 438-468, 2021. DOI: 10.19132/1808-5245274.438-468. Disponível em: <https://seer.ufrgs.br/EmQuestao/article/view/116729>. Acesso em: 10 jun. 2023.

BOOIJ, Geert Evert. **The grammar of words**: an introduction to linguistic morphology. New York: Oxford University Press, 2005.

BRASCHER, Marisa; CAFÉ, Lígia. Organização da informação ou organização do conhecimento? *In*: Encontro Nacional de Pesquisa e Pós-Graduação em Ciência da Informação, 9., 2008, São Paulo. **[Anais eletrônicos]**. São Paulo: ANCIB, 2009. p.

1-14. Disponível em: <https://brapci.inf.br/index.php/res/v/176535>. Acesso em: 31 out. 2023.

BUCKLAND, Michael. Documentation, Information Science, and Library Science in the U.S.A. **Information Processing & Management**, [s. l.], v. 32, n. 1, p. 63-76, 1996. Disponível: [https://doi.org/10.1016/0306-4573\(95\)00050-Q](https://doi.org/10.1016/0306-4573(95)00050-Q). Acesso em: 15 jul. 2023.

BUCKLAND, Michael. Information as thing. **Journal of the American Society for Information Science**, [s. l.], v. 45, n. 5, p. 351-360, 1991. Disponível em: [https://doi.org/10.1002/\(SICI\)1097-4571\(199106\)42:5<351::AID-ASI5>3.0.CO;2-3](https://doi.org/10.1002/(SICI)1097-4571(199106)42:5<351::AID-ASI5>3.0.CO;2-3). Acesso em 10 jul. 2023.

CHARAUDEAU, Patrick; MAINGUENEAU, Dominique. **Dictionnaire d'analyse du discours**. Paris: Seuil, 2002.

COX, Richard J; YAKEL, Elizabeth; WALLACE, David; BASTIAN, Jeannette Allis; MARSHALL, Jennifer. Archival education in North American Library and Information Science schools. **The Library Quarterly**, [s. l.], v. 71, n. 2, p. 141-194, abr. 2001. DOI: 10.1086/603260. Disponível em: <https://www.journals.uchicago.edu/doi/epdf/10.1086/603260>. Acesso em: 31 out. 2023.

CRISTINA AGANETTE, Elisângela; TEIXEIRA, Livia Marangon Duffles; PINTO AGANETTE, Karina de Jesus. A representação descritiva nas perspectivas do século XXI: um estudo evolutivo dos modelos conceituais. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v. 22, n. 50, p. 176-187, set./dez. 2017. DOI: 10.5007/1518-2924.2017v22n50p176. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2017v22n50p176>. Acesso em: 20 dez. 2023.

DOBRESKI, Brian. **Anglo-American library cataloging**. Versão 1.0. [S. l.]: ISKO, 2020. Disponível em: <https://www.isko.org/cyclo/cataloging.htm>. Acesso em: 12 dez. 2023.

EVERT, Stephanie. Corpora and collocations. *In*: LÜDELING, Anke; KYTÖ, Merja. **Corpus linguistics: an international handbook**. Berlin: Mouton de Gruyter, 2009. Disponível em: <https://doi.org/10.1515/9783110213881.2>. Acesso em: 23 jul. 2023.

FUJITA, Mariângela Spotti Lopes. A representação documentária no processo de indexação com o modelo de leitura documentária para textos científicos e livros: uma abordagem cognitiva com protocolo verbal. **PontodeAcesso**, Salvador, v. 7, n. 1, p. 42-66, abr. 2013. Disponível em: <https://periodicos.ufba.br/index.php/revistaici/article/view/8135>. Acesso em: 17 fev. 2024.

GIL, Antônio Carlos. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Atlas, 2008.

GUIMARÃES, José Augusto Chaves. Abordagens teóricas de Tratamento Temático da Informação (TTI): catalogação de assunto, indexação e análise documental. **Ibersid**: revista de sistemas de información y documentación, Saragoça, v. 3, p. 105-117, set. 2009. DOI: 10.54886/ibersid.v3i.3730. Disponível em: <https://www.ibersid.eu/ojs/index.php/ibersid/article/view/3730>. Acesso em: 20 dez. 2023.

GUIMARÃES, José Augusto Chaves. Análise de domínio como perspectiva metodológica em organização da informação. **Ciência da Informação**, Brasília, v. 41, n. 1, p. 13-21, 2014. DOI: 10.18225/ci.inf.v43i1.1415. Disponível em: <https://revista.ibict.br/ciinf/article/view/1415>. Acesso em: 31 out. 2023.

GUIMARÃES, José Augusto Chaves; SANTOS, André Ynada dos; SALES, Rodrigo de; MATOS, Daniela Fernanda de Oliveira. A dimensão conceitual da organização do conhecimento nos congressos da NASKO: análise de conteúdo Bardiana. *In*: CONGRESSO BRASILEIRO DE ORGANIZAÇÃO DO CONHECIMENTO, 3., 2015, Marília. **Organização do Conhecimento e Diversidade Cultural**. Marília: ISKO-Brasil; FUNDEPE, 2015. p. 55-72. Disponível em: <https://isko.org.br/wp-content/uploads/2021/05/Proceedings-ISKO-Brasil-2015.pdf>. Acesso em: 3 dez. 2023.

HJØRLAND, Birger. Concept theory. **Journal of the American Society for Information Science and Technology**, New York, v. 60, n. 8, p. 1519-1536, ago. 2009. DOI: 10.1002/asi.21082. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/asi.21082>. Acesso em: 31 out. 2023.

HJØRLAND, Birger. Domain analysis. **Knowledge Organization**, [s. l.], v. 44, n. 6, p. 436-464, 2017. Disponível em: <https://doi.org/10.5771/0943-7444-2017-6-436>. Acesso em: 10 jun. 2023.

HJØRLAND, Birger. Domain analysis in information science: eleven approaches – traditional as well as innovative. **Journal of Documentation**, [s. l.], v. 58, n. 4, p. 422-462, ago. 2002. DOI: 10.1108/00220410210431136. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/00220410210431136/full/html>. Acesso em: 10 jun. 2023.

HJØRLAND, Birger. Domain analysis: a socio-cognitive orientation for Information Science Research. **Bulletin of the American Society for Information Science and Technology**, [s. l.], v. 30, n. 3, p. 12-28, fev./mar. 2004. DOI: 10.1002/bult.312. Disponível em: <https://asistdl.onlinelibrary.wiley.com/doi/10.1002/bult.312>. Acesso em: 10 jun. 2023.

HJØRLAND, Birger. Theories are Knowledge Organizing Systems (KOS). **Knowledge Organization**, [s. l.], v. 42, n. 2, p. 113-128, 2015. DOI: 10.5771/0943-7444-2015-2-113. Disponível em: <https://www.nomos-elibrary.de/10.5771/0943-7444-2015-2-113/theories-are-knowled-ge-organizing-systems-kos-jahrgang-42-2015-heft-2?page=1>. Acesso em: 31 jul. 2023.

HJØRLAND, Birger. What is Knowledge Organization (KO)? **Knowledge Organization**, [s. l.], v. 35, n. 2/3, p. 86-101, 2008. DOI: 10.5771/0943-7444-2008-2-3-86. Disponível em: <https://www.nomos-elibrary.de/10.5771/0943-7444-2008-2-3-86/what-is-knowledge-organization-ko-jahrgang-35-2008-heft-2-3?page=1>. Acesso em: 29 jul. 2023.

HJØRLAND, Birger; ALBRECHTSEN, Hanne. Toward a new horizon in Information Science: domain-analysis. **Journal of the Association for Information Science and Technology**, New York, v. 46, n. 6, p. 400-425, jul. 1995. DOI: 10.1002/(SICI)1097-4571(199507)46:6<400::AID-ASI2>3.0.CO;2-Y. Disponível em: [https://asistdl.onlinelibrary.wiley.com/doi/10.1002/\(SICI\)1097-4571\(199507\)46:6%3C400::AID-ASI2%3E3.0.CO;2-Y](https://asistdl.onlinelibrary.wiley.com/doi/10.1002/(SICI)1097-4571(199507)46:6%3C400::AID-ASI2%3E3.0.CO;2-Y). Acesso: 10 jun. 2023.

HODGE, Gail. **Systems of Knowledge Organization for digital libraries**: beyond traditional authority files. Washington: DLF, 2000. Disponível em: <https://www.clir.org/pubs/reports/pub91/>. Acesso em: 31 jul. 2023.

INTERNATIONAL SOCIETY FOR KNOWLEDGE ORGANIZATION (ISKO). **ISKO chapters**. [S. l.]: ISKO, 2023a. Disponível em: <http://www.isko.org/chapters.html>. Acesso em: 3 dez. 2023.

INTERNATIONAL SOCIETY FOR KNOWLEDGE ORGANIZATION (ISKO). **ISKO's mission**. [S. l.]: ISKO, 2023b. Disponível em: <https://www.isko.org/about>. Acesso em: 3 dez. 2023.

INTERNATIONAL SOCIETY FOR KNOWLEDGE ORGANIZATION CANADA/UNITED STATES CHAPTER (ISKO-C/US). **About ISKO C-US**. [Canada]: ISKO C-US, c2009-2010a. Página web. Disponível em: <https://iskocus.org/aboutISKO-CUS.php>. Acesso em: 10 jun. 2023.

INTERNATIONAL SOCIETY FOR KNOWLEDGE ORGANIZATION CANADA/UNITED STATES CHAPTER (ISKO-C/US). **Events**. [Canada]: ISKO C-US, c2009-2010b. Página web. Disponível em: <https://iskocus.org/events.php>. Acesso em: 10 jun. 2023.

ISKO News 13. **Knowledge Organization**, [s. l.], v. 20, n. 2, p. 91-93, 1993. DOI: 10.5771/0943-7444-1993-2-91. Disponível em: <https://www.nomos-elibrary.de/10.5771/0943-7444-1993-2-91/isko-news-13-jahrgang-20-1993-heft-2?page=1>. Acesso em: 10 jun. 2023.

LENNON, John; MCCARTNEY, Paul. **I am the walrus**. London: Lennon-McCartney, 1967. Faixa que integra o EP duplo "Magical Mystery Tour" da banda The Beatles. Disponível em: <https://www.thebeatles.com/i-am-walrus>. Acesso em: 10 de jan. 2024.

MARCONI, Marina de Andrade; LAKATOS, Eva Maria. **Fundamentos da metodologia científica**. 5. ed. São Paulo: Atlas, 2003.

MCKIM, Courtney. The value of mixed methods research: a mixed methods study. **Journal of Mixed Methods Research**, [s. l.], v. 11, n. 2, p. 202-222, 2017. DOI:

10.1177/155868981560709. Disponível em:
<https://journals.sagepub.com/doi/10.1177/1558689815607096>. Acesso em: 21 jul. 2023.

ORLANDI, Eni Puccinelli. Michel Pêcheux e a Análise de Discurso. **Estudos da Língua(gem)**, Vitória da Conquista, v. 1, n. 1, p. 9-13, 2005. DOI: 10.22481/el.v1i1.973. Disponível em:
<https://periodicos2.uesb.br/index.php/estudosdalinguagem/article/view/973>. Acesso em: 30 jun. 2023.

PANCHYSHYN, Roman; LAMBERT, Frank; MCCUTCHEON, Sevim. Resource Description and Access adoption and implementation in public libraries in the United States. **Library Resources & Technical Services**, [s. l.], v. 63, n. 2, p. 119-130, apr. 2019. DOI: 10.5860/lrts.63n2.119. Disponível em:
<https://journals.ala.org/index.php/lrts/article/view/6737>. Acesso em: 23 nov. 2023.

PÊCHEUX, Michel. **O discurso**: estrutura ou acontecimento. 2. ed. Campinas: Pontes, 1997.

PROISL, Thomas. **The cooccurrence of linguistic structures**. Erlanger: FAU University Press, 2019. *E-book*. Disponível em:
<https://doi.org/10.25593/978-3-96147-201-7>. Acesso em: 31 out. 2023.

QIN, Jian. Evolving paradigms of knowledge representation and organization: a comparative study of classification, XML/DTD, and ontology. *In*: INTERNATIONAL SOCIETY FOR KNOWLEDGE ORGANIZATION CONFERENCE, 7., 2002, Granada. **Proceedings**. Würzburg: Ergon, 2002. Disponível em:
https://jianqin.metadataetc.org/wp-content/uploads/2019/04/qin_isko2002.pdf. Acesso em: 31 jan. 2023.

QIN, Jian. Knowledge Organization and Representation under the AI Lens. **Journal of Data and Information Science**, [s. l.], v. 5, n. 1, p. 3-17, 2020. DOI: 10.2478/jdis-2020-0002. Disponível em:
<https://sciendo.com/article/10.2478/jdis-2020-0002>. Acesso em: 31 dez. 2023.

ROSENBACH, Anette. Emerging variation: determiner genitives and noun modifiers in english. **English Language and Linguistics**, Cambridge, v. 11, n. 1, p. 143-189, 2007. DOI: 10.1017/S1360674306002140. Disponível em:
<https://www.cambridge.org/core/journals/english-language-and-linguistics/article/abs/emerging-variation-determiner-genitives-and-noun-modifiers-in-english/41F7AA8661A14D44A2AE58ED717DD228>. Acesso em: 31 out. 2023.

RUBI, Milena Polsinelli. Os princípios da política de indexação na análise de assunto para catalogação: especificidade, exaustividade, revocação e precisão na perspectiva dos catalogadores e usuários. *In*: FUJITA, Mariângela Spotti Lopes (org.). **A indexação de livros**: a percepção de catalogadores e usuários de bibliotecas universitárias. São Paulo: Cultura Acadêmica, 2009. p. 81-93. Disponível em: <https://books.scielo.org/id/wcvbc/pdf/bocato-9788579830150-06.pdf>. Acesso em: 31 dez. 2023.

SARDINHA, Tony Berber. Corpus size. **The Specialist**, São Paulo, v. 23, n. 2, p. 103-122, 2003. Disponível em: <https://revistas.pucsp.br/index.php/esp/article/view/9381/6952>. Acesso em: 31 out. 2023.

SARDINHA, Tony Berber. **Linguística de corpus**. Barueri: Manole, 2004.

SAUSSURE, Ferdinand de. **Curso de linguística geral**. São Paulo: Editora Cultrix, 2006.

SINCLAIR, John. **Corpus, concordance, collocation**. Oxford: Oxford University Press, 1991.

SKETCH ENGINE. **enTenTen**: corpus of the English Web. Lexical Computing CZ s.r.o.: [Brun], [202-?]a. Disponível em: <https://www.sketchengine.eu/ententen-english-corpus/#bibliography>. Acesso em: 12 dez. 2023.

SKETCH ENGINE. **Sketch Engine user guide**. Lexical Computing CZ s.r.o.: [Brun], [202-?]b. Disponível em: <https://www.sketchengine.eu/guide/>. Acesso em: 10 jun. 2023.

SKETCH ENGINE. **What can Sketch Engine do with...?** Lexical Computing CZ s.r.o.: [Brun], [202-?]c. Disponível em: <https://www.sketchengine.eu/what-can-sketch-engine-do/#top>. Acesso em: 10 jun. 2023.

SMIRAGLIA, Richard P. A glimpse at Knowledge Organization in North America. **Knowledge Organization**, [s. l.], v. 34, n. 2, p. 69-71, 2007. DOI: 10.5771/0943-7444-2007-2-69. Disponível em: <https://www.nomos-elibrary.de/10.5771/0943-7444-2007-2-69/a-glimpse-at-knowledge-organization-in-north-america-jahrgang-34-2007-heft-2?page=1>. Acesso em: 28 out. 2023.

SMIRAGLIA, Richard P. Chapter 6: epistemology of domain analysis. *In*: SMIRAGLIA, Richard P.; LEE, Hur-Li (org.). **Cultural frames of knowledge**. Milwaukee: Ergon Verlag, 2012. p. 111-124.

SMIRAGLIA, Richard P. **Domain analysis for Knowledge Organization**: tools for ontology extraction. Waltham, MA (USA); Kidlington (UK): Elsevier, 2015.

SOCIEDADE BRASILEIRA PARA ORGANIZAÇÃO DO CONHECIMENTO (ISKO-BRASIL). **ISKO internacional**. [Brasil]: ISKO-Brasil, [202-?]. Disponível em: <https://isko.org.br/isko-internacional/#>. Acesso em: 25 nov. 2023.

SZOSTAK, Rick; OHLY, Peter. **The International Society for Knowledge Organization (ISKO)**. [S. l.]: ISKO, 2020. Disponível em: <https://www.isko.org/cyclo/isko>. Acesso em: 3 dez. 2023.

TENNIS, Joseph. Com o que uma análise de domínio se parece no tocante a sua forma, função e gênero? **Brazilian Journal of Information Science: Research Trends**, Marília, v. 6, n. 1, p. 3-14, jan./jun. 2012. DOI: 10.36311/1981-1640.2012.v6n1.02.p3. Disponível em: <https://revistas.marilia.unesp.br/index.php/bjis/article/view/3026>. Acesso em: 10 jun. 2023.

TENNIS, Joseph. Two axes of domain for domain analysis. **Knowledge Organization**, [s. l.], v. 30, n. 3/4, p. 191-195, 2003. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2879034. Acesso em: 9 jun. 2023.

THE FOUNDING of the International Society for Knowledge Organization Frankfurt, 22 July 1989. **Knowledge Organization**, [s. l.], v. 16, n. 2, p. 71-72, 1989. DOI: 10.5771/0943-7444-1989-2-71. Disponível em: <https://www.nomos-elibrary.de/10.5771/0943-7444-1989-2-71/the-founding-of-the-international-society-for-knowledge-organization-frankfurt-july-22-1989-jahrgang-16-1989-heft-2?page=1>. Acesso em: 3 dez. 2023.

VENANCIO, Renato Pinto. Ser e não ser: as relações históricas entre arquivologia e ciência da informação. **Brazilian Journal of Information Science: Research Trends**, Marília, v. 11, n. 4, p. 23-32, 2017. DOI: 10.36311/1981-1640.2017.v11n4.05.p23. Acesso em: <https://revistas.marilia.unesp.br/index.php/bjis/article/view/7500>. Acesso em: 20 out. 2023.

ZENG, Marcia Lei. Knowledge Organization Systems (KOS). **Knowledge Organization**, [s. l.], v. 35, n. 2-3, p. 160-182, 2008. DOI: 10.5771/0943-7444-2008-2-3-160. Disponível em: <https://www.nomos-elibrary.de/10.5771/0943-7444-2008-2-3-160/knowledge-organization-systems-kos-jahrgang-35-2008-heft-2-3?page=1>. Acesso em: 31 jul. 2023.

APÊNDICE A - LISTA COMPLETA DA WORDLIST

Ordem	Palavra	Tradução	Frequência absoluta
1º	<i>knowledge</i>	conhecimento	3,701
2º	<i>classification</i>	classificação	3,398
3º	<i>information</i>	informação	2,587
4º	<i>library</i>	biblioteca	2,518
5º	<i>organization</i>	organização	2,472
6º	<i>subject</i>	assunto	2,148
7º	<i>term</i>	termo	1,627
8º	<i>system</i>	sistema	1,605
9º	<i>analysis</i>	análise	1,359
10º	<i>category</i>	categoria	1,090
11º	<i>concept</i>	conceito	1,086
12º	<i>user</i>	usuário	1,032
13º	<i>relationship</i>	relação	1,004
14º	<i>domain</i>	domínio	951
15º	<i>class</i>	classe	882
16º	<i>tag</i>	etiqueta; marcação	832
17º	<i>facet</i>	faceta	828
18º	<i>model</i>	modelo	808
19º	<i>language</i>	linguagem	816
20º	<i>process</i>	processo	808
21º	<i>web</i>	rede (<i>internet</i>)	718
22º	<i>structure</i>	estrutura	708
23º	<i>datum</i>	dado	700
24º	<i>ko</i>	OC	692
25º	<i>indexing</i>	indexação	688
26º	<i>cataloge</i>	catalogação	687
27º	<i>represent</i>	representação	667
28º	<i>vocabulary</i>	vocabulário	609
29º	<i>document</i>	documento	575

30°	<i>bibliographic</i>	bibliográfico	547
31°	<i>ontology</i>	ontologia	543
32°	<i>representation</i>	representação	540
33°	<i>content</i>	conteúdo	536
34°	<i>metadata</i>	metadado	466
35°	<i>semantic</i>	semântica	466
36°	<i>heading</i>	cabeçalho	457
37°	<i>object</i>	objeto	456
38°	<i>retrieve</i>	recuperar	453
39°	<i>entity</i>	entidade	451
40°	<i>standard</i>	padrão	447
41°	<i>retrieval</i>	recuperação	436
42°	<i>scheme</i>	esquema	400
43°	<i>ddc</i>	CDD	391
44°	<i>description</i>	descrição	378
45°	<i>relation</i>	relação	372
46°	<i>catalog</i>	catálogo	356
47°	<i>taxonomy</i>	taxonomia	354
48°	<i>catalogue</i>	catálogo	298
49°	<i>lcsb</i>	LCSH	274
50°	<i>organize</i>	organizar	268
51°	<i>koss</i>	SOCs	236
52°	<i>conceptual</i>	conceitual	249
53°	<i>librarian</i>	bibliotecário	241
54°	<i>index</i>	indexar	232
55°	<i>decimal</i>	decimal	221
56°	<i>thesaurus</i>	tesauro	215
57°	<i>systems</i>	sistemas	214
58°	<i>frbr</i>	FRBR	211
59°	<i>authority</i>	autoridade	204
60°	<i>terminology</i>	terminologia	197

61°	archival	arquivo	195
62°	descriptive	descritivo	189
63°	bibliographical	bibliográfico	198
64°	documentation	documentação	188
65°	hierarchy	hierarquia	178
66°	hierarchical	hierárquico	169
67°	rda	RDA	152
68°	schema	esquema	150
69°	classificatory	classificatório	149
70°	dictionary	dicionário	146
71°	cataloger	catalogador	145
72°	descriptor	descritor	144
73°	controlled	controlado	140
74°	archives	arquivos	130
75°	indexer	indexador	126
76°	rdf	RDF	124
77°	subclass	subclasse	121
78°	headings	cabeçalhos	116
79°	udc	CDU	102
80°	lcc	LCC	90
81°	bibframe	BIBFRAME	81
82°	foaf	FOAF	76
83°	kdc	KDC	74
84°	bcc	BCC	72
85°	marc	MARC	69
86°	aar	AAT	67
87°	librarianship	biblioteconomia	66
88°	archivist	arquivista	66
89°	ifla	IFLA	64
90°	aacr2	AACR2	61
91°	archive	arquivo	60

92°	classificationist	classificacionista	60
93°	classifier	classificador	53
94°	owl	OWL	53
95°	terminological	terminológico	40
96°	organisation	organização	35
97°	representative	representativo	33

APÊNDICE B - LISTA COMPLETA DAS KEYWORDS (SINGLE-WORDS)

Ordem	Palavra	Tradução	Frequência absoluta		Total
			NASKO	Referência	
1º	<i>cataloge</i>	catalogação	687	83,511	389.7
2º	<i>ddc</i>	CDD	391	24,746	373.1
3º	<i>lcsh</i>	Cabeçalhos de assunto da Biblioteca do Congresso	274	3,357	347.9
4º	<i>koss</i>	SOCs	263	7,601	313.4
5º	<i>bibliographic</i>	bibliográfico	547	83,094	311.3
6º	<i>indexing</i>	indexação	688	144,957	274.2
7º	<i>frbr</i>	Requisitos Funcionais para Registros Bibliográficos	211	4,444	263.7
8º	<i>classification</i>	classificação	3,398	1,150,480	230.6
9º	<i>ko</i>	OC	692	214,882	206.0
10º	<i>classificatory</i>	classificatório	149	3,588	188.9
11º	<i>ontology</i>	ontologia	543	206,819	166.6
12º	<i>bibliographical</i>	bibliográfico	189	33,170	164.7
13º	<i>cataloger</i>	catalogador	145	12,450	161.9
14º	<i>kos</i>	SOC	292	93,566	155.1
15º	<i>facet</i>	faceta	828	379,290	154.6
16º	<i>thesaurus</i>	tesauro	215	57,731	148.7
17º	<i>taxonomy</i>	taxonomia	354	169,603	126.2
18º	<i>indexer</i>	indexador	126	22,052	124.6
19º	<i>headings</i>	cabeçalhos	116	16,626	122.7
20º	<i>rda</i>	RDA	152	45,520	117.3
21º	<i>heading</i>	cabeçalho	457	285,778	108.4
22º	<i>neurodiversity</i>	neurodiversidade	91	8,377	107.8
23º	<i>udc</i>	CDU	102	17,196	107.2
24º	<i>bibframe</i>	Bibliographic Framework Initiative	81	1,463	106.6
25º	<i>lis</i>	biblioteconomia	147	52,424	106.6
26º	<i>retrieval</i>	recuperação	436	275,793	106.4
27º	<i>intellection</i>	intelecção	80	5,185	99.4

28°	semantic	semântica	466	326,102	99.0
29°	foaf	Friend of a Friend Ontology 2000	76	6,227	93.1
30°	metatheory	metateoria	67	1,716	88.0
31°	kdc	Korean Decimal Classification	74	8,595	87.6
32°	classificationist	classificationista	60	63	81.0
33°	vocabulary	vocabulário	609	565,621	79.9
34°	aacr2	AACR2	61	3,280	78.3
35°	instantiation	instanciação	81	32,415	71.5
36°	epistemological	epistemológico	101	56,342	70.9
37°	metadata	metadado	466	479,383	70.9
38°	categorial	categorial	55	3,035	70.9
39°	ifla	Federação Internacional de Associações e Instituições Bibliotecárias	64	19,155	65.9
40°	lcc	Library of Congress Classification	90	52,266	65.5
41°	rdf	Resource Description Framework	124	98,526	64.1
42°	epistemology	epistemologia	107	77,275	63.8
43°	pragmatism	pragmatismo	92	58,465	63.5
44°	aat	Art & Architecture Thesaurus	67	26,677	63.1
45°	epistemic	epistêmico	85	51,913	62.1
46°	decimal	decimal	221	233,480	61.8
47°	librarianship	biblioteconomia	66	30,110	59.9
48°	citeulike	CiteULike	44	979	58.8
49°	lod	Linked Open Data	71	39,217	58.5
50°	rmv	Répertoire de vedettes-matière	47	5,535	58.5
51°	ontological	ontológico	94	73,727	57.6
52°	notational	notacional	47	7,389	56.9
53°	subclass	subclasse	121	115,321	56.6
54°	authorial	autoral	56	22,395	55.6

55°	categorization	categorização	94	79,769	55.1
56°	co-citation	co-citação	41	1,852	54.1
57°	descriptor	descriptor	144	169,591	51.5
58°	bcc	Basic Concepts Classification	72	54,642	51.5
59°	actants	agentes	39	2,124	51.3

APÊNDICE C - LISTA COMPLETA DAS KEYWORDS (MULTI-WORDS TERMS)

Ordem	Palavra	Tradução	Frequência absoluta		Total
			NASKO	Referência	
1º	knowledge organization	organização do conhecimento	1,681	4,449	2,094.0
2º	subject headings	cabeçalhos de assunto	277	15,964	294.5
3º	information science	ciência da informação	401	65,723	259.5
4º	controlled vocabulary	vocabulário controlado	185	10,882	210.8
5º	facet analysis	análise de facetas	154	324	205.5
6º	classification scheme	esquema de classificação	172	14,760	186.0
7º	decimal classification	classificação decimal	139	1,480	182.2
8º	knowledge organization system	sistema de organização do conhecimento	133	994	175.7
9º	domain analysis	análise de domínio	135	2,745	173.5
10º	subject indexing	indexação de assunto	126	966	166.6
11º	library classification	classificação de biblioteca	111	918	147.0
12º	bibliographic classification	classificação bibliográfica	107	218	143.3
13º	information retrieval	recuperação de informação	170	38,894	139.7
14º	classification system	sistema de classificação	165	41,398	132.3
15º	indexing language	linguagem de indexação	98	421	130.9
16º	faceted classification	classificação facetada	95	886	126.0
17º	subject analysis	análise de assunto	89	738	118.4
18º	classification theory	classificação (teoria)	84	403	112.4
19º	dewey decimal	decimal de dewey	86	3,146	110.2
20º	metadata schema	esquema de metadados	83	3,059	106.5
21º	bibliographic record	registro bibliográfico	87	8,236	103.3
22º	information organization	organização da informação	75	2,951	96.5

23°	subject head	cabeçalho de assunto	78	5,634	96.3
24°	knowledge strategy	estratégia do conhecimento	71	718	94.7
25°	knowledge representation	representação do conhecimento	85	13,113	94.4
26°	category system	sistema de categorias	71	1,979	92.8
27°	resource description	descrição do recurso	76	6,448	92.8
28°	evolutionary order	ordem evolutiva	68	206	91.5
29°	systematic indexing	indexação sistemática	64	0	86.4
30°	semantic relation	relação semântica	67	2,894	86.4
31°	sears list	Sears List	64	79	86.3
32°	archival science	arquivologia	62	1,236	82.1
33°	classification designer	designer de classificação	60	0	81.1
34°	knowledge organizations systems	sistemas de organização do conhecimento	59	261	79.4
35°	social tag	tag social	61	2,504	79.2
36°	library trend	tendência da biblioteca	60	1,853	78.7
37°	subject representation	representação de assunto	57	151	76.9
38°	general subject	assunto geral	61	5,869	75.3
39°	information seek	busca de informação	69	15,512	74.4
40°	knowledge structure	estrutura do conhecimento	58	3,632	74.1
41°	subject classification	classificação de assunto	55	2,246	71.8
42°	dewey decimal classification	classificação decimal de dewey	52	301	70.1
43°	digital library	biblioteca digital	100	58,160	69.2
44°	cataloging practice	prática de catalogação	51	430	68.6
45°	library catalog	catálogo de biblioteca	64	16,224	68.4
46°	subject access	acesso ao assunto	54	4,351	68.3
47°	dictionary catalog	catálogo dicionário	50	135	67.6
48°	literary warrant	garantia literária	49	305	66.1

49°	congress classification	classificação do congresso	49	1,272	65.1
50°	subject catalogue	catálogo de assunto	48	379	64.7
51°	main class	classe principal	55	9,355	64.6
52°	bibliographical classification	classificação bibliográfica	47	0	63.8
53°	author keyword	palavras-chave do autor	47	1,729	62.0
54°	information professional	profissional da informação	54	11,708	61.4
55°	material type	tipo de material	57	15,851	61.3
56°	web facet	facetadas da web	45	0	61.1
57°	professional indexer	profissional indexador	45	258	60.8
58°	class number	número da classe	49	5,832	60.7
59°	discourse analysis	análise de discurso	56	16,721	59.6
60°	associative relationship	relação associativa	44	438	59.3
61°	linked datum	Liked Data	54	15,224	58.6
62°	discovery system	sistema de descoberta	45	2,702	58.5
63°	genre tag	tag de gênero	42	866	56.3
64°	library science	biblioteconomia	54	18,656	56.1
65°	metadata element	elemento de metadado	43	3,271	55.5
66°	title keyword	palavras-chave de título	41	426	55.4
67°	social discovery	descoberta social	41	805	55.0
68°	functional requirement	Requisitos Funcionais	51	15,972	54.9
69°	universe of knowledge	universo do conhecimento	40	394	54.1
70°	knowledge discovery	descoberta de conhecimento	48	12,723	53.9
71°	information seeking behavior	comportamento de busca de informação	40	938	53.6
72°	ko system	Sistema de Organização do Conhecimento	39	62	53.0
73°	concept theory	teoria do conceito	39	197	52.9

74°	<i>american library association</i>	Associação Americana de Bibliotecas	51	19,183	52.7
75°	<i>information object</i>	objeto de informação	41	3,900	52.4
76°	<i>classificatory change</i>	mudança bibliográfica	37	0	50.4
77°	<i>bibliographic records</i>	registros bibliográficos	37	755	49.8
78°	<i>information resource</i>	recursos de informação	60	39,629	49.4
79°	<i>classification research</i>	pesquisa de classificação	36	347	48.8
80°	<i>epistemic community</i>	comunidade epistêmica	35	1,442	46.6
81°	<i>authority record</i>	registro de autoridade	36	3,478	46.4