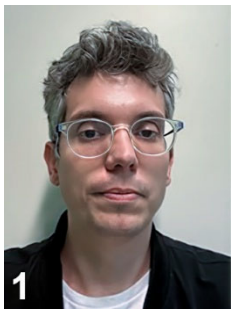
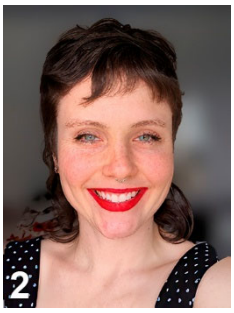
Knowledge Organization Systems (KOS) in the Context of ISKO: A Domain Analysis of the Brazilian and North American Chapters

Thiago Henrique Bragato Barros*, Maria Amália Cassol Lied**, Waleska da Silva Rocha***, Kamila de Andrade Moura****

* ** *** ****Universidade Federal do Rio Grande do Sul (UFRGS),
Rua Ramiro Barcelos, 2705 – Santana, Porto Alegre
* bragato.barros@ufrgs.br, ** mariamalialied@gmail.com,
*** waleskadasilvarocha@gmail.com, **** kamoura15@gmail.com





Thiago Henrique Bragato Barros is an Associate Professor in the Department of Information Science at the Federal University of Rio Grande do Sul (UFRGS) and a permanent Professor in the Graduate Programs in Information Science at UFRGS and UFSC. Editor in Chief of the Journal Em Questão. President of the International Society for Knowledge Organization (2022-2026). Managing Editor of the journal Knowledge Organization. Coordinator of the Graduate Program in Information Science – PPGCIN UFRGS (2020-2024). Led a Working Group at Ancib. Chaired ISKO Brazil. Coordinated the Undergraduate Course in Archival Studies and Directed the College of Archival Studies. Organized national and international events in the field of information. Representative of the Information area on the Advisory Committee at FAPERGS (2022-2024). Has experience in archival studies and knowledge organization.

Maria Amália Lied has a Master's degree in Literature, specializing in Comparative Literature, from the Federal University of Pelotas (2019). Bachelor's degree in Languages, emphasizing Portuguese Language and Portuguese Literature, from the Federal University of Rio Grande do Sul (2014). Bachelor's degree in Library Science from the Federal University of Rio Grande do Sul (2024). Served as a Scientific Initiation Scholarship recipient (CAPES-2022/FAPERGS 2022-2024) and as a member of the research group Organization and Representation of Knowledge: Linguistic Approaches in Archives and Libraries – ORCALAB.

Waleska da Silva Rocha is a Master's student in the Postgraduate Program in Information Science at the Federal University of Rio Grande do Sul (PPGCIN – UFRGS), with an undergraduate degree in Library Science from the Federal University of Rio Grande do Sul (UFRGS). Her current research uses domain and discourse analysis to analyze Brazilian researchers' scientific output in knowledge organizations within regional chapters of the International Society for Knowledge Organization (ISKO). She is a member of the research group Organization and Representation of Knowledge: Linguistic Approaches in Archives and Libraries (ORCALAB).

Kamila Moura is a librarian with a degree from the Federal University of Rio Grande do Sul (UFRGS) and a Master's degree in Information Science from the Postgraduate Program in Information Science (PPGCIN/UFRGS), registered with the Regional Library Council (CRB/10) under number 2528. Affiliated with the research group Organization and Representation of Knowledge: Linguistic Approaches in Archives and Libraries (ORCALAB/UFRGS) and the Legal Information and Documentation Group of Rio Grande do Sul (GIDJ-RS). Experienced in legal information, implementation and use of document management and metadata software, application of ABNT and ISO standards,

cataloging processes, classification (DDC and UDC), and indexing. My research interests include knowledge organization, knowledge organization systems, information retrieval, terminology, information technology, knowledge management, diplomatics, and domain analysis.

Barros, Thiago Henrique Bragato, Maria Amália Cassol Lied, Waleska da Silva Rocha, and Kamila de Andrade Moura. 2024. "Knowledge Organization Systems (KOS) in the Context of ISKO: a Domain Analysis of the Brazilian and North American Chapters". *Knowledge Organization* 51, no. 8: 582-599. 16 references. DOI:10.5771/0943-7444-2024-8-582.

Abstract: Objective: It aims to map, analyze thematically, semantically, and discursively the articles published in the area of Knowledge Organization within the five published volumes of the events of the International Society of Knowledge Organization Brazil (ISKO-Brasil) and the eight volumes of the North American Symposium on Knowledge Organization (NASKO). Methodology: The methodology used was to survey pre-established pivot statements of the scientific production published in the annals of both events with the help of Sketch Engine software as a tool and approach 6 – Historical studies of structures and services of information in domains, 8 – Epistemological and critical studies of different paradigms, assumptions and interests in domains, 10 – Studies of structures and institutions of scientific and professional communication in a domain, from Hjørland (2017), and 13 – Discourse analysis of domain analysis as a methodological contribution, from Smiraglia (2015) and Barros (2023). Results: The survey yielded 131 articles for ISKO-Brasil and 132 documents for NASKO. It was found that, even though the two corpora analyzed are within ISKO's scope, there are divergences regarding the understanding of concepts, as well as their relationship with the epistemological discussion of the area and convergences concerning the concepts of 1) Domain Analysis; 2) Organization Systems; 3) Concept Theory; 4) Classification Systems. Conclusion: The analysis made it possible to envision Knowledge Organization as a theoretical and applied area based on Concept Theory and Domain Analysis.

Received: 15 October 2024; Accepted 29 November 2024.

Keywords: ISKO; Domain analysis; Knowledge Organization Systems; Sketch Engine.

1.0 Introduction

This work seeks to identify and compare the discourse communities within the field of Knowledge Organization (KO) across the chapters of the International Society for Knowledge Organization (ISKO), based on an analysis of the proceedings from the following events: the Brazilian Conference on Knowledge Organization and Representation (ISKO-Brazil) (2011-2019) and the North American Symposium on Knowledge Organization (NASKO) (2007-2021). Thus, this research aims to map and analyze thematically, semantically, and discursively the articles published in the field of Knowledge Organization within the five volumes published from the ISKO-Brazil events and the eight volumes of NASKO.

To achieve this, the research is supported by the works of Barros and Laipelt (2021) and Oliveira et al. (2022), as they serve as guiding references for studies of this nature. The NASKO Conference first took place in 2007, totaling eight editions since then. The Brazilian Conference on Knowledge Organization began four years later, with a total of six events to date. Both conferences remain regular and are held approximately every two years.

Given the relevance of these events within the ISKO framework and the active contribution of their communities to the scientific production in the field, these discourse communities provide a basis for analyzing how their activities resonate in the conceptualizations and perceptions of KO and Knowledge Organization Systems (KOS).

It is worth noting that this work is part of a more extensive study funded by CNPq that aims to analyze trends in the Literature on Knowledge Organization Systems (KOS) within the ISKO context as a whole.

2.0 Knowledge Organization

Knowledge Organization (KO) aims to provide access to information in all fields of human understanding or activity, both for those within and outside a given field. Knowledge is organized, after all, so that it can be used (even a scientist in a laboratory organizes their data in some order to refer back to it). From this statement of purpose, we can derive several theses about how the field should operate.

With this in mind, we can understand KO as a field associated with Information Science that is dedicated to studying processes of representation and systematization of concepts, producing as outputs Knowledge Organization Systems (KOS) – ontologies, taxonomies, controlled vocabularies, thesauri, among others. Through these systems, KO structures concepts and their relationships within a knowledge domain, building models for the representation and organization of information (Brascher and Café 2010; Barros and Laipelt 2021, 40). Given its importance, some of the pivotal terms selected for this analysis are KOS, namely ontologies, thesauri, and taxonomies. Their application varies across different social contexts, supported by various technological foundations that enable Knowledge Organization through management and access (Barros 2021, 64).

According to Barros (2021, 62), analysis through discourse does not strictly limit terms to a conceptual perspective – as is traditionally approached in KO – but, as a "theoretical framework," enables historical, social, and ideological analysis. From this perspective, it is understood that the documents examined in this study go beyond the technical necessity of observing the terminological development of these discourse communities, also addressing how their conceptual trends are articulated socially and epistemologically.

Since the delimitation and understanding of the subject significantly influence the progression of the research, the events and pivotal terms were outlined to understand how conceptions regarding representation and KOS evolve within the Brazilian and North American events. For this purpose, the Sketch Engine software was used as a tool, along with approaches 6, 8, 10, and 13 of domain analysis as methodological frameworks.

3.0 Domain Analysis

Domain analysis is an approach initially formulated by Hjørland and Albrechtsen (1995). It emphasizes the importance of studying knowledge domains through their language, structure, and communities, as these share unique theories, terminologies, and paradigms. It involves analyzing the operation of Information Science (IS) from a social perspective, focusing on the contexts in which information circulates. Specialized domains offer fertile ground for understanding how these dynamics occur (Lopez-Huertas, 2015; Smiraglia, 2015).

Hjørland (2002; 2017) outlined 11 approaches for applying domain analysis, providing broad analytical coverage within the context of IS. For the purposes of this study, the following approaches are used:

Approach 6 – Historical studies of information structures and services within domains, aiming better to understand the domain, its structure, and organization.
Approach 8 – Epistemological and critical studies of different paradigms, assumptions, and interests within domains to map the epistemological construction of the area.
Approach 10 – Studies of the structures and institutions of scientific and professional communication within a domain to understand the informational cycle configuration of the analyzed domain.
Approach 13 (based on Smiraglia, 2015 and Barros, 2023) – Discourse analysis to comprehend institutionalized discourse within the domain.

These approaches provide a multifaceted framework for exploring how knowledge organization unfolds conceptually and socially within specialized fields.

4.0 Sketch Engine

Sketch Engine is online software for linguistic analysis of texts. It utilizes textual corpora (language samples) to allow researchers to identify what is representative, rare, and obsolete within a corpus.

For this study, the following system features were used:

- Word Sketch (analyzing grammatical and collocational behavior of words).
- Word Sketch Difference (comparing contrasting collocations).
- Keywords (for terminological extraction).
- Wordlist (for frequency analysis).

The search was performed using lemmas, following the standard definition by Booij (2005, 3), who describes a lemma (or lexeme) as a word viewed in its "abstract sense," encompassing its various morphological forms. These variations, termed word forms, represent its "concrete sense." For instance, the concrete words "walks," "walked," "walking," can all be classified as forms of the lexeme [walk] (Booij 2005, 3).

A standard selection of the first ten terms from each material was used, as it would have been impractical to present all terms from the lists and graphs generated by the tool. If the term's influence within each corpus was minimal, the selection was reduced.

Considering that the two discourse communities analyzed – ISKO-Brazil and NASKO – operate in different languages, this distinction was critical in the search and analysis of terms in Sketch Engine. ISKO-Brazil primarily uses Brazilian Portuguese (the focus of this study, though there are also articles in English and Spanish). At the same time, NASKO predominantly features American English speakers (with some British English variations observed). Given that each community displayed distinct grammatical behaviors, the same search strategy could not be uniformly applied.

Word Sketch: Summarizes the collocational and grammatical behavior of a word, showing its relationships with other words through various categories termed grammatical relations (Sketch Engine, 2023).

Word Sketch Difference (WSD): Compares two lemmas to analyze their collocations and grammatical relations (Sketch Engine, 2023).

Keywords: Extracts key terms either as single words (individual terms) or multi-word terms (phrases). These terms typically define or characterize the content or topic of a corpus, document, or text (Sketch Engine, 2023).

Wordlist: This tool generates frequency lists for various types of terms and provides metrics on how often specific terms appear in the corpus (Sketch Engine, 2023).

Given that Sketch Engine is a technical tool for corpus analysis, the researcher must interpret the data qualitatively. For this purpose, the Concordance feature was used. This tool provides the context in which words appear, offering greater reliability to terms that align with the scope of Knowledge Organization (KO).

Sketch Engine allows for the generation of charts and data tables. The study aims to use these outputs to compare, organize, and analyze information extracted from both textual corpora to employ domain analysis effectively. This methodological integration combines quantitative and qualitative approaches to enhance the depth and accuracy of the findings.

5.0 Results

A preliminary survey of pre-established statements in the scientific production of the International Society of Knowledge Organization – Brazil (ISKO-Brazil) was conducted, based on the proceedings of the Brazilian Conference on Knowledge Organization and Representation (2011-2019) and the North American Symposium on Knowledge Organization (NASKO) (2007-2021). This resulted in analyzing 131 articles from ISKO-Brazil and 132 documents (127 articles and 5 extended abstracts) from NASKO. For the analysis of the Brazilian chapter, the articles were divided into three dimensions: epistemological, applied, social, cultural, and political. However, this study focused only on the applied and social, cultural, and political dimensions, using exclusively articles published in Portuguese.

The statements used for selecting the ISKO-Brazil corpus were "Indexação" (Indexing), "Sistemas de Organização do Conhecimento" (Knowledge Organization Systems), "Representação Documental" (Document Representation), "Representação do Conhecimento" (Knowledge Representation), "Representação da Informação" (Information Representation), "Representação e Organização do Conhecimento" (Representation and Organization of Knowledge), "Taxonomia" (Taxonomy), "Tesauro" (Thesaurus), and "Ontologia" (Ontology). For the North American chapter, the respective terms were "Knowledge Organization Systems," "Knowledge Organization," "Taxonomies," "Ontologies," and "Thesaurus."

After the initial analysis and selection of the research corpus, the Sketch Engine software was used to examine the publications from these two discourse communities. The ISKO-Brazil proceedings, comprising five volumes, resulted in a corpus of 532,496 words in Portuguese. The NASKO proceedings, comprising eight volumes, constituted a corpus of 583,674 words in English. The difference and plurality of terms in the Brazilian chapter are due to the specific characteristics of its community, which exhibits more excellent terminological dispersion compared to its North American counterpart.

5.1 ISKO-BRASIL

The data had to be cleaned to generate the Wordlist for the ISKO-Brazil corpus, as the software detects the frequency of the entire corpus. Therefore, punctuation marks, prepositions, articles, connectors, and similar elements were disregarded, focusing solely on terms that are significant concepts within the corpus. This process ensures that only meaningful terms relevant to the research objectives are included, as illustrated in Table 1.

	Word	Frequency
1°	informação	3,535
2°	conhecimento	3,339
3°	organização	2,645
4°	representação	1,687
5°	indexação	1,502
6°	análise	1,338
7°	acesso	1,199
8°	ciência	1,163
9°	pesquisa	1,113
10°	termos	1,064

Table 1. Wordlist ISKO-Brasil.

Some words in the Wordlist originate from headers and footers in the conference proceedings, such as the names of the events. This is the case for the first four words on the list. However, despite this origin, these words hold value for the corpus since they are also part of the statements used to form the corpus. Beyond these four high-frequency words, the fifth word, "indexação" (indexing), highlights an important aspect of the corpus. Unlike the previous terms, it is not part of the conference names and does not appear in headers or footers. This indicates a significant interest in research on the topic of indexing within ISKO-Brazil, which is crucial for Knowledge and Information Organization and Representation.

Data cleaning was also performed on the tables developed using Keywords. For multi-word terms, to ensure the list was accurate and representative of the corpus, terms originating from event headers and footers, frequently recurring journal names, and English-language terms were excluded. This refinement aimed to ensure that only relevant terms closely aligned with the research focus were included, as shown in Table 2.

	Words
1º	política de indexação
2º	sistemas de organização
3º	análise de domínio
4º	análise de assunto
5º	descrição arquivística
6º	linguagem de indexação
7º	tratamento temático
8º	linguagens documentárias
9º	processo de indexação
10º	indexação de imagens

Table 2. Keywords (multi-words terms) ISKO-Brasil.

The 10 keywords identified in the ISKO-Brazil corpus most pertain to topics related to thematic representation, focusing on issues surrounding indexing, subject analysis, or thematic treatment. In Archival Science, archival description stands out as a fundamental process in organizing and representing archival knowledge. Other notable terms in this context include those related to the processes that organize knowledge, such as knowledge organization systems and documentary languages.

Regarding methodological procedures, domain analysis is a widely used theoretical framework within ISKO-Brazil, highlighting its relevance in the field.

Data cleaning was also performed in Table 3, which addresses single words. Words in English and author names were excluded, focusing solely on terms with significant conceptual importance for the corpus. This refinement ensures that the single-word terms selected are meaningful and relevant to the research objectives.

	Words
1º	indexação
2º	tesauros
3º	ontologia
4º	arquivístico
5º	tesauro
6º	isko
7º	informacional
8º	terminológico
9º	folksonomia
10º	cdd

Table 3. Keywords (single-words) ISKO-Brasil.

In Table 3, some words are part of the statements used to define the corpus, such as indexação (indexing), tesauros (thesauri), tesauro (thesaurus), and ontologia (ontology). It is worth noting that there appears to be an error in the software's identification process, as it treated the singular (tesauro) and plural (tesauros) forms as separate words. Ideally, the software should have grouped these terms regardless of their grammatical number or gender.

Regarding the terms themselves, similar to the multi-word terms, indexação ranks first, highlighting a significant interest among the ISKO-Brazil epistemic community in issues related to the indexing of materials. Another noteworthy aspect of this list is the presence of terms related to knowledge organization systems (KOS), such as tesauro (thesaurus), ontologia (ontology), and folksonomia (folksonomy), alongside a classification system widely used in Brazilian libraries, the Dewey Decimal Classification (CDD). This underscores the importance of KOS for the ISKO-Brazil discourse community.

In the context of the Word Sketch tool, the analysis of the ISKO-Brazil corpus focused on combining terms with adjectives (term + adjective) and verbs (verb + term). Based on insights from the Wordlist and Keywords tools, the study narrowed its scope to the following terms: organização (organization), conhecimento (knowledge), representação (representation), informação (information), and indexação (indexing). This targeted approach allows a deeper exploration of how these concepts interact within the corpus.

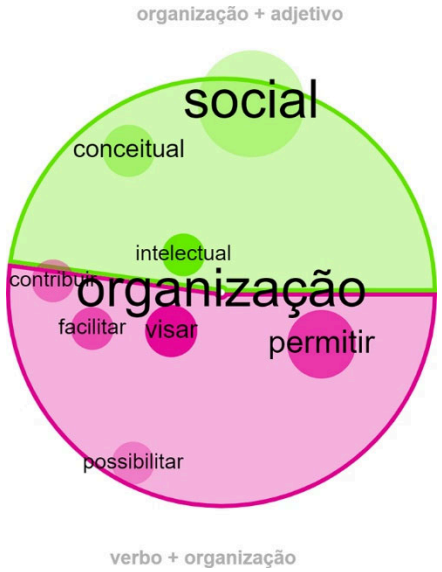


Figure 1. Word Sketch "organization".

In Figure 1, analyzing the grammatical relationship between the term organização (organization) and verbs, the closest and most frequently related verbs are permitir (to enable) and visar (to aim). This suggests a focus within the community on facilitating and aiming for improved organization – whether of knowledge or information.

However, when examining the term organização linked to adjectives, it is notable that terms included in this research's statements are not retrieved. This is due to the software's inability to recognize the term as part of a prepositional phrase. When using such filters, errors occur in data generation.

An important point to highlight is the association of the term organização with the adjective social (social). This connection aligns with Hjørland's (2008) observations on the social organization of knowledge, emphasizing how knowledge is organized within different domains, reflecting the realities and structures of those contexts. This underscores the interplay between social dynamics and the organization of knowledge and information.

Regarding the term conhecimento (knowledge), as shown in Figure 2, the adjective most closely associated with it is responsável (responsible). However, this case highlights another instance where the term originates from the title of the 5th edition of the event, *Responsible Knowledge Organization: Promoting Democratic and Inclusive Societies*, appearing on every page of the articles from that year's proceedings.

As for the verbs connected to the term *conhecimento*, they reflect core activities studied in the field. Examples include *representar* (to represent), *organizar* (to organize), *produzir* (to produce), and *compartilhar* (to share). These verbs encapsulate fundamental processes and goals within the area of Knowledge Organization, underscoring the community's focus on how knowledge is created, structured, and disseminated.

Regarding the term representação (representation) in Figure 3, the adjective most closely associated with it is temático (thematic). Using the Concordance tool, it is evident that thematic representation is a recurring theme in several articles within the corpus. This finding aligns with earlier analyses using the Wordlist and Keywords tools, highlighting the community's strong focus on indexing and thematic treatment.

Other noteworthy adjectives linked to representação include documental (documental), bibliográfico (bibliographic), arquivístico (archival), imagético (imagistic), and gráfico (graphic). These terms emphasize specific materials requiring equally specialized forms of representation, reflecting the diversity and specificity of representation challenges in different contexts.

Figure 2. Word Sketch "knowledge".

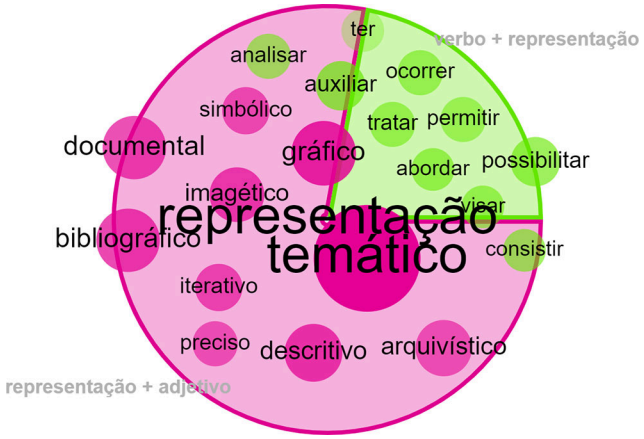


Figure 3. Word Sketch "representation".

Verbs associated with representação emerge in patterns similar to those seen with the term organização. Verbs like permitir (to enable) and visar (to aim) appear prominently alongside others like possibility (to make possible). These verbs reflect actions crucial to enabling and facilitating the representation of knowledge and/or information, highlighting their relevance within the discourse of Knowledge Organization.

Regarding the term informação (information) in Figure 4, the adjective most closely associated with it is arquivístico (archival), forming the term informação arquivística (archival information). Another related term, informação orgânica (organic information), also emerges, reflecting a trend within the ISKO-Brazil discourse community toward topics connected to Archival Science.

As for the verbs associated with informação, they align with typical activities in the field, including representar (to represent), organizar (to organize), recuperar (to retrieve), and buscar (to search), among others. These verbs underscore the central processes of managing and utilizing information, reflecting the community's focus on core Information Science practices *arquivística* (archival science). Another related term, *informação orgânica*(organic information), also emerges, reflecting a trend within the ISKO-Brazil discourse community toward topics connected to Archival Science.

As for the verbs associated with *informação*, they align with typical activities in the field, including *representar* (to represent), *organizar* (to organize), *recuperar* (to retrieve), and *buscar* (to search), among others. These verbs underscore the central processes of managing and utilizing information, reflecting the community's focus on core Information Science practices.

As previously mentioned, the term indexação (indexing) was selected because it appears with high frequency in the corpus articles. This trend was evident from the initial analysis stages during the corpus formation. Using the Word Sketch tool (see Figure 5), it was observed that there is a strong focus on *automatic indexing*, indicating advancements in indexing processes; *multimodal indexing*, which is closely linked to ontologies; and *social indexing*, which is connected to folksonomies and the earlier discussion on the social organization of knowledge within the communities where knowledge is indexed. Additionally, the mention of *specific indexing languages* aligns with these themes, emphasizing tailored approaches to indexing.

To address the limitation of the Word Sketch tool, which in Portuguese cannot analyze terms as part of prepositional phrases due to software errors, the Word Sketch Difference(WSD) tool was utilized. This alternative allowed for a more nuanced analysis of the relationships and contexts involving the term *indexação*, providing insights into its connections and usage within the corpus.

In Figure 6, an analysis was conducted on the terms *organização* (organization) and *representação* (representation), focusing on their usage in conjunction with *de* followed by a noun. The findings reveal a community tendency to discuss *conhecimento* (knowledge) and *informação* (information), with a stronger emphasis on *organização do conhecimento*(knowledge organization) and *representação da informação* (information representation). While the term *informação* is centrally positioned in the graph, this indicates a balanced use of the concepts *representação da informação* and *organização da informação*.

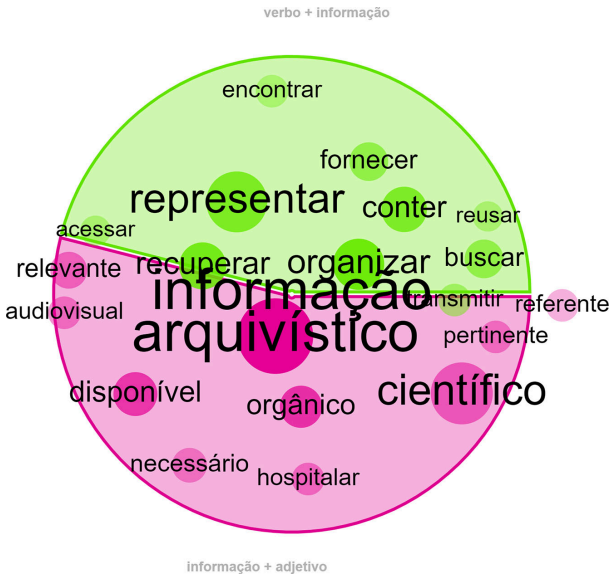


Figure 4. Word Sketch "information".

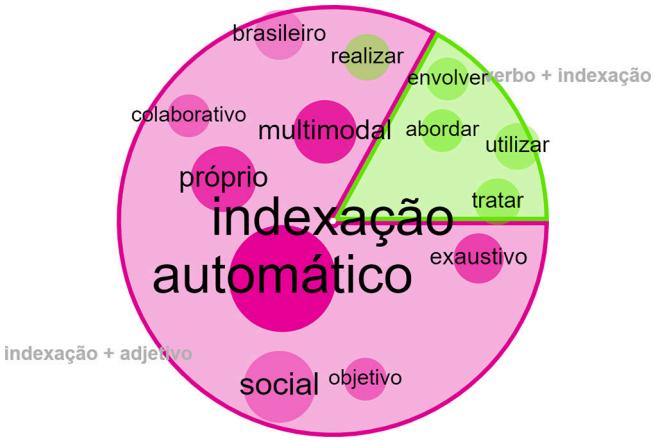


Figure 5. Word Sketch "indexing".

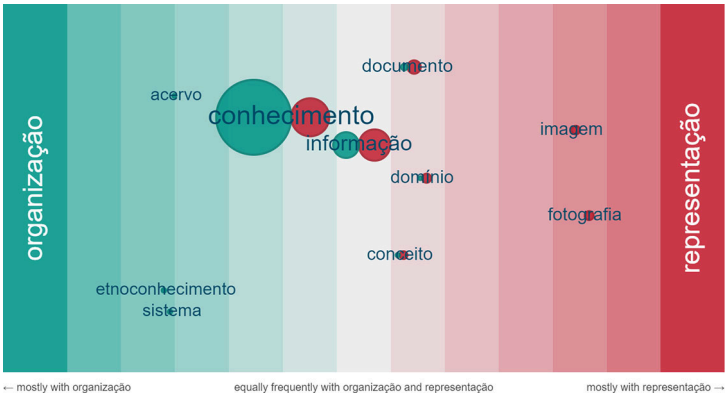


Figure 6. WSD (KNOWLEDGE-REPRESENTATION).

Additionally, referencing Brascher and Café's (2008) insights into these concepts, it is evident that *organização do conhecimento* focuses on organizing concepts, while *representação da informação* is concerned with recording information, and that's a consensus among the Brazilian community. This distinction highlights the complementary nature of these practices within Knowledge Organization and Representation for the analyzed community.

To support these findings on *organização* and *representação*, an analysis was conducted using the Word Sketch Difference tool for the terms *conhecimento* (knowledge) and *informação* (information), applying the filter *... de conhecimento/informação*, forming prepositional phrases. The results in Figure 7 further clarify the nuanced relationship and distinct focus areas of these terms within the discourse, reinforcing the conceptual balance and specificity observed in the corpus.

Similar to the previous graph, the same conclusions can be drawn. There is a clear inclination toward *organização do conhecimento* (knowledge organization), *representação da informação* (information representation), and *organização da informação* (information organization). A notable aspect of this graph is the proximity of the word *recuperação* (retrieval) to the term *informação* (information), suggesting that the ISKO-Brazil community is also focused on issues related to *recuperação da informação* (information retrieval).

In addition to these terms, both *conhecimento* (knowledge) and *informação* (information) are also connected – though less prominently – to *produção* (production) and *gestão* (management). This highlights the community's concern with developing and managing knowledge and information.

For the analysis of Knowledge Organization Systems (KOS), specifically ontologies (ontology) and taxonomies (taxonomy), the Word Sketch Difference tool was used with the filter e ou (and), as depicted in Figure 8. This approach aimed to explore the relationships and co-occurrence patterns of these two KOS concepts, shedding light on their comparative and complementary usage within the discourse community.

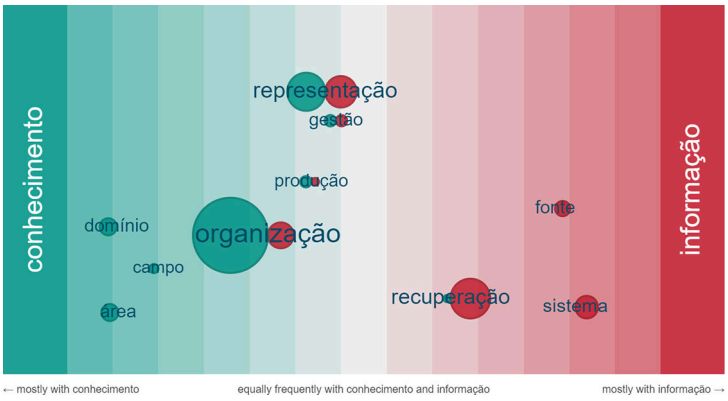


Figure 7. WSD (KNOWLEDGE-INFORMATION).

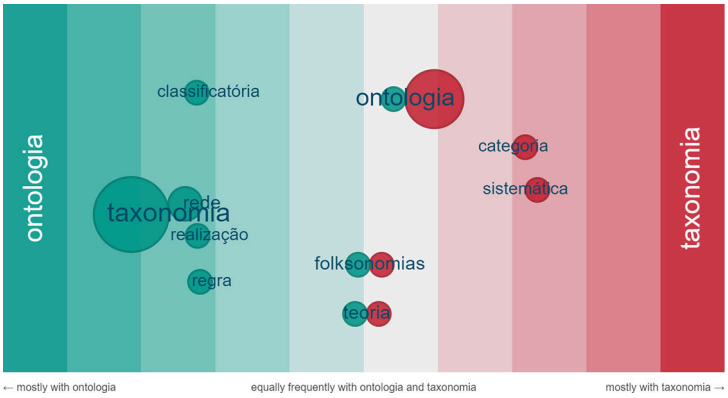


Figure 8. WSD (Ontology-Taxonomy).

The terms *ontologia* (ontology) and *taxonomia* (taxonomy) hold significant importance within the corpus and the broader field of Knowledge Organization. Even before the corpus analysis was conducted using specialized software, it was apparent that more studies and discussions focused on ontologies than *taxonomies*. This predominance can be attributed to ontologies' central role in modern Knowledge Organization Systems (KOS), particularly in areas that require sophisticated semantic structures and frameworks for managing complex relationships between concepts. Ontologies are often seen as pivotal tools in advanced information systems, enabling semantic interoperability and enhancing data integration, which makes them a critical focus of research and application. Taxonomies, while essential for hierarchical organization and classification, seem to receive relatively less attention, perhaps due to their perceived simplicity compared to the rich conceptual frameworks provided by ontologies.

Another critical observation is that these terms often appear grouped with other KOS in the texts, such as in the phrase "ontologies, taxonomies, and thesauri." This grouping underscores the interconnectedness of these systems and their complementary roles in the organization and retrieval of knowledge. While the term *tesauro* (thesaurus) did not appear prominently in this analysis, the emergence of *folksonomia* (folksonomy) – a newer form of KOS – was notable. This inclusion reflects the community's awareness and exploration of contemporary, user-driven systems alongside more traditional, expert-curated ones. Using the Concordance tool revealed that terms like *ontologia*, *taxonomia*, and *folksonomia* often appear in contexts where KOS are listed as examples, illustrating the diversity of systems used in organizing knowledge. This reinforces the idea that the ISKO-Brazil community actively engages with a broad spectrum of KOS, exploring both their theoretical underpinnings and practical applications. The dynamic interplay between these systems reflects ongoing efforts to address the evolving needs of information systems, demonstrating a balance between traditional methodologies and innovative approaches.

5.2 NASKO

The Wordlist analysis found that the term *knowledge* is the most frequently occurring word in the NASKO corpus,

with a total of 3,701 occurrences. This highlights the centrality of the concept of knowledge within the discussions and publications of the NASKO community. The following knowledge, as shown in Table 4, includes other high-frequency terms such as *classification*, *information*, *library*, and *organization*.

The prominence of these terms reflects the thematic focus of the corpus on key aspects of Knowledge Organization, particularly the processes and systems used to classify and organize information within libraries and other information environments. The occurrence of *libraries* further emphasizes the field's solid historical and practical connection to library and information science. These terms illustrate the primary subjects of interest and point to the interdisciplinary nature of the research, encompassing theoretical discussions about knowledge and classification, as well as applied aspects related to the management and organization of information resources. This distribution of terms suggests a balance in the NASKO community's focus, combining conceptual exploration with practical implications in information science.

	Word	Frequency
1º	knowledge	3,701
2º	classification	3,398
3º	information	2,587
4º	library	2,518
5º	organization	2,472
6º	term	1,627
7º	system	1,605
8º	analysis	1,359
9º	category	1,090
10º	concept	1,086

Table 4. Wordlist NASKO.

It is possible to conclude that these are highly recurrent words within the thematic scope of this work. They fit within a terminological pattern as they encompass terms commonly used by discourse communities in Knowledge Organization. In constructing this table, symbols, irrelevant words (such as those appearing only in headers and not representative of the corpus), prepositions, and author names were disregarded. This data cleaning was also applied to the tables developed from the Keywords analysis, as the objective was to thematically examine the relationships and recurrence of terms within the corpus.

Table 5 shows the lemma related to cataloging activity in the first position. An analysis using the Concordance and Word Sketch tools reveals that the topic is often addressed regarding its practice and standardization.

	Word
1º	cataloge
2º	DDC
3º	LCSH
4º	KOS
5º	bibliographic
6º	FRBR
7º	indexing
8º	classification
9º	ko
10º	ontology

Table 5. Keywords (single-word) NASKO.

The selected corpus includes representation tools widely developed and explored in the North American context, such as the Dewey Decimal Classification (DDC) and the Library of Congress Subject Headings (LCSH). These systems have long been central to organizing and accessing library collections in the region. The British conceptual model Functional Requirements for Bibliographic Records (FRBR) was also identified. Notably, while these terms are included in their full form, their acronyms – such as DDC, LCSH, and FRBR – are more frequently used within the texts, except Knowledge Organization (KO), often referenced explicitly. This prominence of general classifications versus specialized classifications, a characteristic of the community

The term ontology stands out among the proposed statements with a higher frequency of other Knowledge Organization Systems (KOS) included in the statements. This prominence reflects the increasing importance of ontologies within the field, particularly for their role in structuring complex relationships and enhancing semantic understanding across diverse contexts.

When it comes to terms comprising multiple words, known as multi-word terms, the analysis reveals, as shown in Table 6, that the leading term is *knowledge organization*. This reinforces the thematic centrality of KO as a cornerstone concept in the discourse community, reflecting the focus on organizing knowledge both as a theoretical framework and as a practical field of study. The prominence of this term aligns with the broader objectives of Knowledge Organization, underscoring its interdisciplinary relevance and its role in advancing the management and representation of information across various domains.

	Word
1º	knowledge organization
2º	subject headings
3º	information science
4º	controlled vocabulary
5º	facet analysis
6º	classification scheme
7º	decimal classification
8º	knowledge organization system
9º	domain analysis
10º	library classification

Table 6. Keywords (multi-words terms) NASKO.

The third keyword of the analysis, the discipline of Information Science, is identified, reinforcing the connection between the broader field and its subfield, Knowledge Organization (KO). The presence of the term Knowledge Organization Systems (KOS) further underscores NASKO's interest in this subject area. Within the established scope, two terms are classified as KOS under the Classification and Categorization group: subject headings and classification schemes (Zeng 2008, 161).

Terms such as facet analysis, decimal classification, subject heading, and library classification highlight the community's predominant focus on representation. Additionally, domain analysis, a methodology frequently studied and applied within KO, is a crucial theme, ranking ninth in prominence within the corpus. This highlights the ongoing relevance of domain analysis in exploring epistemological and practical aspects of knowledge organization.

The Word Sketch tool was applied to the terms knowledge, organization, and system to complement this understanding. These terms were chosen because they align with and contrast with the pivotal statements proposed, such as knowledge organization and knowledge organization systems. This analysis aimed to uncover their relationships with other terms in the corpus. Additionally, specific KOS, such as taxonomy, thesaurus, and ontology, were examined for their distinct roles and connections within the discourse.

The strongest relationship observed for the term knowledge is with organization, emphasizing its frequent association with the field's name, knowledge organization. The term system also demonstrates a strong connection, aligning with the pivotal statement of knowledge organization systems. This reflects the NASKO community's thematic focus on the theoretical and applied aspects of organizing knowledge through systems and frameworks.

Figure 9 shows that beyond knowledge organization, there is significant interest in related concepts such as its representation, structure, and organizational strategies. These findings highlight the multifaceted nature of KO within the discourse, showing an interplay between theoretical constructs and practical applications. They further the understanding of how knowledge is structured, represented, and operationalized within the field.

Regarding terms that indicate actions related to the word, specifically verbs, Figure 9 reveals a predominance of similar collocations such as to organize, to represent, and to create. These verbs reflect the core processes within the field of Knowledge Organization. Additionally, other verbs like to share, produce, acquire, and store stand out, as they are associated with the basic operations involved in managing and processing information and knowledge. These actions form the foundation of how knowledge is handled, from acquisition to storage and dissemination.

The term organization, as one that qualifies or connects with other terms, shows a particularly strong relationship with the word *sistema* (system), as observed in Figure 10. This connection underscores the focus on systems as a central concept in organizing knowledge, reflecting their role in structuring and operationalizing the processes of Knowledge Organization. The interplay between *organization* and *system* is indicative of the community's emphasis on frameworks and infrastructures that support the organization and accessibility of knowledge and information.

Other terms, such as *practice*, *literature*, and *standardization*, are notable in their proximity to the central conceptual node in the corpus. This proximity highlights their integral role in structuring discussions within the analyzed discourse. The inclusion of the word *community* in the data visualization suggests a pronounced interest in research exploring the dynamics and interactions within knowledge organization communities. This encompasses both the narrower scope of NASKO (North American Symposium on Knowledge Organization) studies and the broader community engaged in the organization of knowledge.

Additionally, the analysis revealed the significant grammatical and thematic relationships between the term *system* and key associated concepts such as *organization*, *information*, *classification*, and *knowledge*. These associations underline the term's foundational role in the domain of knowledge organization, reflecting its usage in phrases like "Knowledge Organization Systems" (KOS) and "Information Retrieval Systems".

Despite the strong semantic linkage between the terms *knowledge* and *information* with the central term *system*, neither appears in immediate proximity to the core of the graph. A deeper exploration through concordance analysis clarified that these terms often occur as parts of multi-word expressions, such as "Systems for Information Retrieval" and "Systems for Knowledge Organization." This highlights their contextualized application within specific systemic frameworks rather than isolated mentions.

Further investigation into the role of the term *system* in these frameworks revealed that verbs such as *evolve*, *organ-*

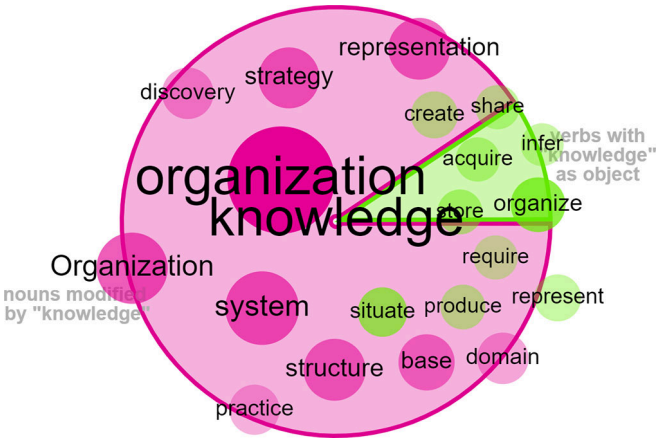


Figure 9. Word Sketch "knowledge".

Figure 10. Word Sketch "organization".

ize, *design*, *create*, *use*, and *base* are frequently associated with it (Figure 11). These verbs describe activities fundamental to the lifecycle of knowledge systems, from conceptual development to practical implementation. Additionally, terms like *facet* signify the influence of faceted approaches, emphasizing modular and dynamic perspectives on system structuring within the domain.

Both *knowledge* and *information* exhibit a strong semantic connection to the central term *system*. However, the graphical representation does not position them close to the center. This spatial arrangement reflects the insights gained through a more in-depth concordance analysis, which revealed that these terms are typically part of multi-word expressions. Examples include "Information Retrieval Systems" and "Knowledge Organization Systems." Within such structures, the terms are conceptually tied to *system*, but they do not appear adjacent to it as standalone terms. This nuanced relationship can be further observed through the term *retrieval*, which exhibits a weaker overall relationship in the graph while showing closer proximity to the central node.

When examining the term *system* as an object, specific verbs emerged as directly associated with the concept of Knowledge Organization Systems (KOS). These verbs include evolve, organize, design, create, use, and base. At their core, these actions relate to the lifecycle of systems, encompassing their conceptualization, development, optimization, and application. To validate these connections, a detailed concordance analysis was conducted, confirming that these verbs frequently appear in contexts directly linked to KOS.

Additionally, the term "facet " surfaced as a key linked concept, indicating that some authors within the knowledge organization community often analyze these systems through a faceted perspective. This suggests an emphasis on modular and multidimensional approaches to system development and implementation, highlighting the flexibility and adaptability inherent in knowledge systems' conceptual

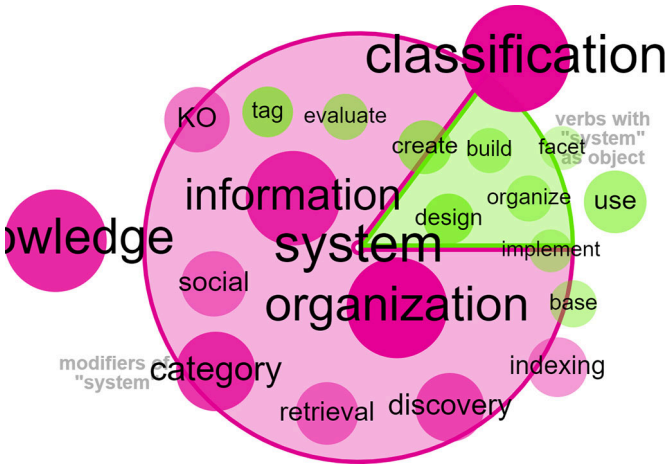


Figure 11. Word Sketch "system".

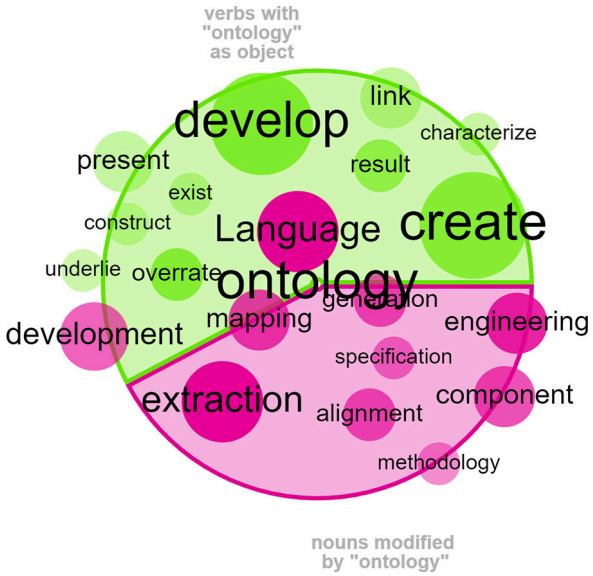


Figure 12. Word Sketch "ontology".

frameworks. This approach aligns with the broader goal of tailoring systems to meet diverse and evolving informational needs.

According to Figure 12, the term *ontology* is predominantly associated with concepts such as *language*, *extraction*, and *engineering*. *This indicates* a strong interdisciplinary connection between Knowledge Organization (KO) and topics within Computer Science. This alignment underscores the knowledge organization community's interest in broadening its research horizons to encompass disciplines fundamental to the development of Knowledge Organization Systems (KOS), particularly ontologies.

Various aspects of ontology are explored within the NASKO discourse, both within its specialization and broader contexts. These include *ontology mapping*, components, *generation*, development processes, and specifications. While these studies are not yet extensive, they highlight a clear research trajectory within this discourse community, reflecting an interest in leveraging ontologies to meet complex organizational and retrieval challenges.

Regarding verbs most closely associated with *ontology*, there is a noticeable emphasis on terms such as *creation* and *development*. This suggests a strong focus on the processes of building and refining ontologies, which are seen as critical tools for structuring and integrating knowledge. These processes reflect the ongoing efforts of the community to enhance the theoretical and practical utility of ontologies in addressing evolving informational and semantic needs.

This interdisciplinary perspective reinforces the pivotal role of ontologies within Knowledge Organizations and reveals a growing synergy with computational methodologies, pointing to a dynamic and expanding research agenda.

Furthermore, as shown in Figure 13, other Knowledge Organization Systems (KOS), such as *taxonomy* and *ontology*, are significantly related to the term *thesaurus*. This relationship will be explored in more depth using the Word Sketch Difference tool. Regarding the verbs in which *thesaurus* functions as an object, the most recurrent ones identified are *structure* and *produce*. These terms align with processes fundamental to creating and organizing such schema, emphasizing the structured and systematic approach to developing thesauri. Among the KOS analyzed, *taxonomy* appears the fewest times across the categories of verbs and adjectives in the Word Sketch analysis. To address this limitation, two specific categories were selected for focused examination, narrowing the analysis to five key terms. These terms represent the most dominant associations within the corpus (Figure 14).

This limited representation highlights that, among the examined KOS, *taxonomy* is the least explored by the knowledge organization community. This relative lack of attention suggests that while *taxonomy* holds a recognized place within the broader framework of knowledge organization, its conceptual development and practical applications may not currently be a priority compared to other systems like *ontology* and *thesaurus*.

This analysis observed that the terms *T1*, *ontology*, *knowledge*, *alignment*, and *approach* were predominantly associated with *taxonomy*. The Concordance tool determined that the strongest correlation, albeit relatively low, is with *ontology*. This suggests that *taxonomy* and *ontology* are occasionally examined together, potentially reflecting complementary or overlapping functions in specific contexts within Knowledge Organization.

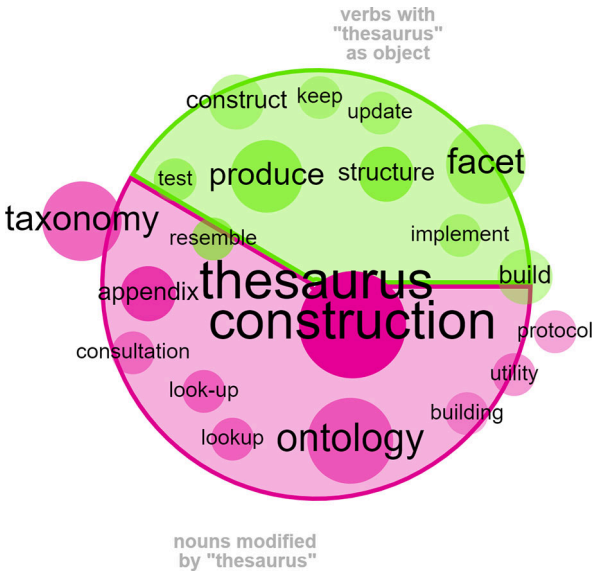


Figure 13. Word Sketch “thesaurus”.

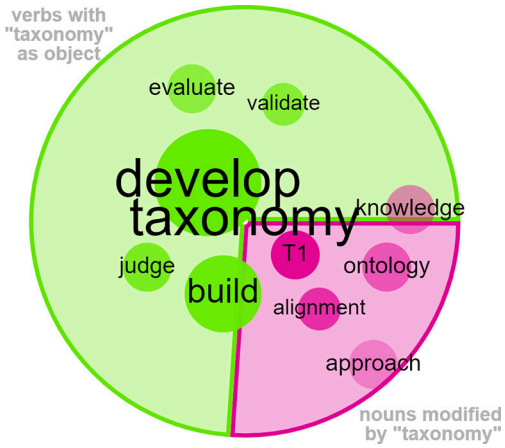


Figure 14. Word Sketch “taxonomy”.

When examining *taxonomy* as the object of verbs, the terms *developed and built* emerged as the most prominent correlating actions. These verbs emphasize the active processes of constructing and refining taxonomies, highlighting their role as tools that require deliberate design and implementation efforts.

Figure 15 further explores the qualifiers related to the terms *knowledge* and *information*. In this analysis, these terms are identified as objects of the grammatical structure, while the qualifiers serve as subjects of the corresponding sentences. This grammatical interplay reflects how *knowledge* and *information* are contextualized and described within the discourse.

The term system displayed a balanced occurrence between the two key concepts analyzed, both in its positioning on the central axis and in the proportion identified across the two spheres. However, there was a stronger relational impact of the term organization with knowledge than with information. Similarly, the concept of representation leaned more toward knowledge than information. Another term positioned at the center of the analysis was object, which showed a closer association with information. This distinction demonstrates that materiality is more closely related to information, often perceived as a tangible entity, whereas knowledge tends to be associated with abstract constructs.

Additional terms such as retrieval and resource also contributed to this analysis, reinforcing the notion of information as a more tangible and actionable concept within the discourse. These terms highlight the functional and pragmatic aspects of working with information, contrasting its application against the more theoretical and conceptual framing of knowledge.

Both knowledge and information act as qualifiers that establish relationships with the terms organization and representation. However, as depicted in Figure 16, these relationships are more strongly inclined toward organization than representation. This suggests that within the analyzed discourse, the structuring and management aspects of organization take precedence over the descriptive and interpretative dimensions of representation.

This distinction reflects the field's practical orientation, emphasizing the organization of knowledge and information as foundational to knowledge systems. Representation serves as a supporting element within this broader framework. Such findings underline the duality of theoretical and applied focus in the study of knowledge organization.

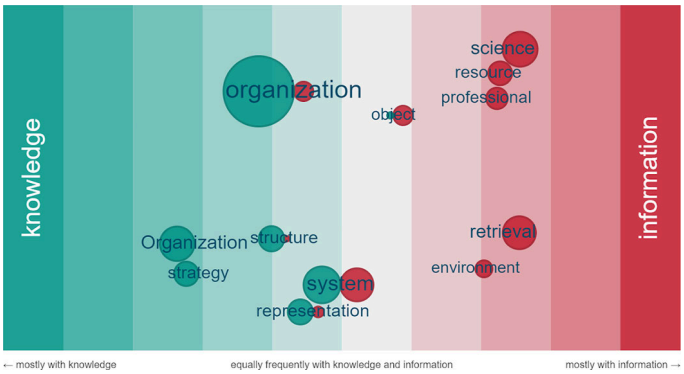


Figure 15. WSD (knowledge-information).

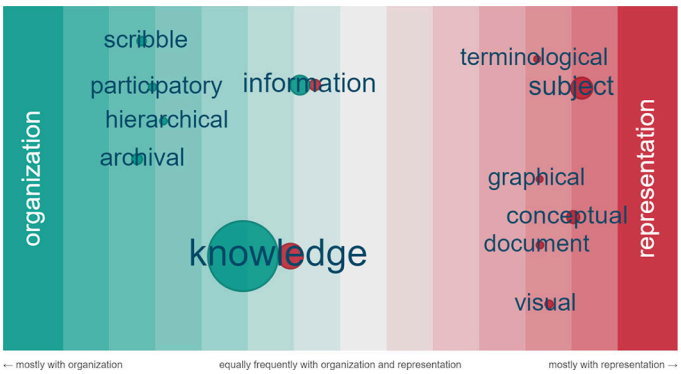


Figure 16. WSD (organization-representation).

The terms *conceptual* and *subject* are qualifiers exclusively associated with *representation*, indicating a strong thematic alignment with the concept of thematic representation. This includes various forms such as conceptual representations (both external and internal), semantic conceptual representations, and subject-based representations. These associations emphasize the role of representation in capturing and organizing thematic and conceptual elements within Knowledge Organization.

In contrast, other qualifiers associated with *organization* include *hierarchical*, *archival*, and *participatory*. These terms reflect different dimensions of organization, addressing its structural (hierarchical), domain-specific (archival), and collaborative (participatory) aspects. These qualifiers are notable for their repeated presence across multiple articles, signifying their importance within the broader discourse of organizational systems.

To better understand the relationship between Knowledge Organization Systems (KOS), a comparative analysis of ontology and taxonomy was conducted (Figure 17). This comparison aims to shed light on the interplay and distinctions between these two systems, which often serve complementary roles within the domain. By focusing on pre-selected pivotal statements, this analysis highlights the nuanced ways in which *ontology* and *taxonomy* are applied and conceptualized within the knowledge organization community.

The findings from this analysis not only delineate the specific contexts in which these terms operate but also reflect broader trends in how Knowledge Organization Systems are integrated into theoretical and practical framework

During the initial phase of the study, which involved manually analyzing each article, a significant pattern of concatenation among Knowledge Organization models was observed. This indicates a tendency within the field to interconnect various Knowledge Organization Systems (KOS) rather than treat them in isolation.

The use of the *and/or* analysis provided greater clarity in visualizing these recurring connections. At the center of the graphical representation, the term *thesaurus* emerged with a stronger relationship to *ontology*. Concordance analysis further revealed that, in most instances where these terms co-occur, their relationship is characterized by *equality*. This pattern was also identified between *taxonomy* and *thesaurus*, indicating that the discourse community does not prioritize highlighting contrasts between these KOS. Instead, it underscores their complementary and integrated roles within the broader framework of knowledge organization.

Another notable finding is the relationship between *ontology* and *folksonomy*, which juxtaposes formal and informal models of knowledge organization. Ontologies are typically formalized systems used in structured contexts, while folksonomies are user-driven, often informal, approaches widely employed in digital and online environments. This contrast reflects the community's recognition of these models' varying applicability based on context, particularly their common association with internet-based applications.

The term *Web*, frequently linked to *ontology*, especially in discussions of the Semantic Web, reinforces this perspective. It highlights ontologies' pivotal role in enabling semantic interoperability and structured knowledge management in online environments, further validating the intertwined relationship between formal ontological models and their application in digital and internet-based systems.

6.0 Conclusion

Based on the analyses of the two chapters, several conclusions can be drawn about the respective discourse communities. Firstly, although both communities operate within the scope of ISKO, an international entity fostering theoretical, methodological, and conceptual collaboration, they have notable differences in their understanding and treatment of concepts, methodologies, and practices. These dif-

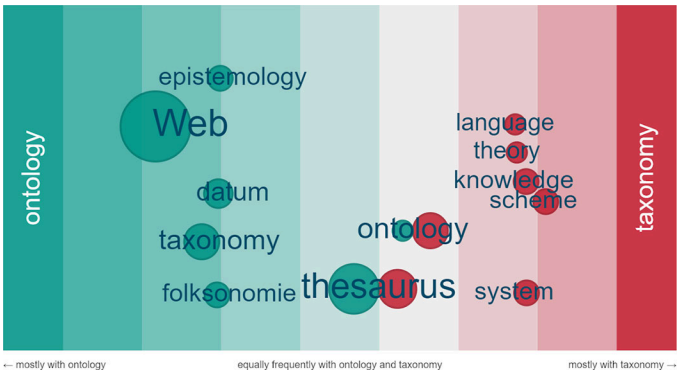


Figure 17. WSD (ontology-taxonomy).

ferences are evidenced by the diverse perspectives and overlapping voices present in the analyzed corpus.

One striking distinction is the prominence of archival themes in the Brazilian chapter, a focus that is not mirrored to the same extent in the North American chapter. Similarly, indexing policies are a prevalent topic within the Brazilian context, whereas they receive comparatively less attention in North American discourse.

Despite these differences, there are significant areas of convergence, particularly in their shared focus on:

1. Domain Analysis: A methodological cornerstone for both communities.
2. Knowledge Organization Systems (KOS): Recognized as essential frameworks for organizing information.
3. Concept Theory: Treated as a foundational theoretical framework.
4. Classification Systems: Highlighted as critical tools for structuring and categorizing knowledge.

These commonalities indicate that the communities operate from a methodological framework grounded in empirical reality. They balance theoretical inquiry with practical application, demonstrating alignment in their reliance on concept theory as a foundational element and domain analysis as a methodological approach.

As this research continues, further chapters will be analyzed to construct a more precise and more comprehensive picture of these communities and their shared and distinct attributes.

Numerous advantages have been observed regarding semantic-discursive analysis as an approach within domain analysis. The use of semantic tools for this purpose has proven to be significant and representative. At this stage, no disadvantages have been identified; on the contrary, this analytical approach has demonstrated its value and versatility. Continued exploration of alternative methods will further enhance the depth and breadth of insights derived from this research.

References

Barros, Thiago Henrique Bragato and Ricardo do Carmo Fonseca Laipeit. 2021. "Uma análise de domínio da área de Organização e Representação do Conhecimento no contexto do periódico *Em Questão*." *Em Questão* 27, no. 4: 438–68. https://doi.org/10.19132/1808-5245274.438-468

Barros, Thiago Henrique Bragato. 2021. "A análise de discurso de Michel Pêcheux e a organização do conhecimento: possibilidades teórico-metodológicas." In *Ensaios em organização do conhecimento*, edited by Isabel dos Santos Amorim, and Rovilson de Sales), 61-78. Florianópolis: UDESC. https://dx.doi.org/10.5965/978-65-88565-20-9

Barros, Thiago Henrique Bragato. 2023. "Michel Pêcheux's Discourse Analysis: An Approach to Domain Analyses". *Knowledge Organization* 50, no. 1: 3-10. https://doi.org/doi.org/10.5771/0943-7444-2023-1-3

Booij, Geert E. 2005. *The Grammar of Words: An Introduction to Linguistic Morphology*. New York: Oxford University Press.

Brascher, Marcia and Luciana Café. 2010. "Organização da informação ou organização do conhecimento?" In *IX Encontro Nacional de Pesquisa e Pós-graduação em Ciência da Informação 28 September-1 October 2010 São Paulo, SP, Brazil.* ENANCIB IX, São Paulo: Escola de Comunicação e Artes da USP. http://hdl.handle.net/20.500.11959/brapci/176535

Hjørland, Birger and Albrechtsen, Hanne. 1995. "Toward a New Horizon in Information Science: Domain Analysis". *Journal of the Association for Information Science and Technology* 46, no. 6: 400–25. DOI: https://doi.org/10.1002/(SICI)1097-4571(199507)46:6<400::AID-ASI2>3.0.CO;2-Y

Hjørland, Birger. 2002. "Domain Analysis in Information Science: Eleven Approaches – Traditional as Well as Innovative". *Journal of Documentation* 58, no. 4: 422–62. https://doi.org/10.1108/00220410210431136

Hjørland, Birger. 2008. "What is Knowledge Organization (KO)?" *Knowledge Organization* 35, nos. 2-3: 86–101. DOI: https://doi.org/10.5771/0943-7444-2008-2-3-86

Hjørland, Birger. 2017. "Domain Analysis." *Knowledge Organization* 44, no. 6: 436–64. DOI: https://doi.org/10.5771/0943-7444-2017-6-436

Hodge, Gail. 2000. *Systems of Knowledge Organization for Digital Libraries: Beyond Traditional Authority Files*. Washington, DC: DLF.

Lexical Computing CZ S.R.O. [202?]. "*Sketch Engine*". Sketch Engine (website). https://www.sketchengine.eu/

López-Huertas, María Jesús. 2015. "Domain Analysis for Interdisciplinary Knowledge Domains". *Knowledge Organization* 42, no. 8: 570–83. https://doi.org/10.5771/0943-7444-2015-8-570

Oliveira, Caliel Cardoso, Mauricio Coelho da Silva, Thiago Henrique Bragato Barros and Ana Maria Mielniczuk de Moura. 2022. "Domain Analysis in Knowledge Organization Research". In *XXII Encontro Nacional de Pesquisa em Ciência da Informação 7-11 November 2022.* ENANCIB XXII. porto Alegre: UFRGS. https://enancib.ancib.org/index.php/enancib/xxiienancib/paper/view/993

Sketch Engine. 2020. "*Sketch Engine User Guide*". *Sketch Engine* (website). https://www.sketchengine.eu/guide/

Smiraglia, Richard P. 2015. *Domain Analysis for Knowledge Organization: Tools for Ontology Extraction*. Waltham, MA: Elsevier.

Zeng, Marcia Lei. 2008. "Knowledge Organization Systems (KOS)". *Knowledge Organization* 35, nos. 2-3: 160–82. DOI: https://doi.org/10.5771/0943-7444-2008-2-3-160