ADVANCES IN KNOWLEDGE ORGANIZATION
Proceedings of International ISKO Conference - ISSN: 0938-5495 - 2026 - São Paulo

Brazilian scientific production in Knowledge Organization: an analysis based on the EIKO categories.

Sônia Oliveira Matos Moutinho

São Paulo State University (UNESP) and Federal Institute of Education, Science and Technology of Piauí (IFPI), Brazil.

Regina Ferreira da Rocha

School of Technology, Teaching Unit of the State Center for Technological Education 'Paula Souza', Brazil.

Mario Barité

University of the Republic, Uruguay.

Abstract

This study analyzes Brazilian scientific production in Knowledge Organization (KO) from 1993 to 2025, based on a bibliometric approach and thematic categorization grounded in the *ISKO Encyclopedia of Knowledge Organization* (IEKO). The corpus consists of 1,000 documents retrieved from BRAPCI, analyzed using quantitative indicators and content analysis. The results indicate a predominance of studies focused on the central concepts of KO, followed by systems (KOS) and processes (KOP), evidencing a strong theoretical orientation. A lower incidence of applied and methodological research is observed. In addition, emerging categories related to artificial intelligence, open science, and the social dimensions of information are identified. It is concluded that the field presents a dual structure, combining traditional cores and emerging themes, which points to the need to update the area's classification structures.

Keywords: Taxonomy. Knowledge Organization – Categorization. Classification.

1.0 Introduction

Knowledge Organization (KO) constitutes a disciplinary field in progressive consolidation on the international scene, reflecting the theoretical, methodological, and applied development of Information Science. In recent decades, there has been a broadening of approaches focused on the representation, structuring, and management of informational content in multiple contexts—scientific, technological, and sociocultural—which has given KO a strategic role in mediating between information, knowledge, and society (Hjorland 2016; Guimarães 2020).

In the Brazilian context, the trajectory of KO follows the institutionalization of Information Science, consolidating itself as one of the most dynamic research areas, both in terms of academic production and in the formulation of instruments and policies for organizing

digital collections and repositories (Alves and Oliveira 2016). This expansion, however, still demands systematic investigations that comprehensively examine the profile, themes, and impact of national scientific production in relation to the global development of the field.

The analysis of scientific production is essential because it makes it possible to identify research trends and gaps, guide public policies and institutional decisions, as well as assess the impact and visibility of national science in the international scenario (Villalobos and Santos 2015). In this sense, mapping Brazilian production in KO is configured as a diagnostic and strategic planning tool for strengthening the area. According to Souza (2013), in the global context, Brazil occupied the eighth position in publications indexed in the *Web of Science* on KO, with 9 records, while countries such as Canada (33 publications), Denmark (24) and France (17) presented substantially higher rates, as shown in Figure 4 of his study. More recently, a bibliometric study conducted on the Scopus information platform observed that Brazil is the second country in scientific production in KO with 10% of the global production in the field (Farías, Aguirre-Ligüera and Barité 2022).

Given this scenario, this study aims to: (i) provide an initial overview of Brazilian scientific production in Knowledge Organization (KO) over the period 1970–2025, based on the BRAPCI database; and (ii) propose a categorization that organizes this production according to the thematic subareas of KO, in order to contribute to understanding its cognitive structure and its position within the international context.

Thus, the following research problem emerges: how does Brazilian scientific production in KO, from 1970 to 2025, reflect the consolidation process and thematic diversity of the field globally?

From this perspective, the following section revisits the historical evolution of KO in Brazilian Information Science, highlighting how the transition from classical collection organization practices in Library Science to the configuration of KO as a structuring subdiscipline and, more recently, as a field of research with relative theoretical and institutional autonomy, constitutes the essential backdrop for understanding the scientific production analyzed in this study.

2.0 Categorization and Classification in KO

As an interdisciplinary field where epistemological, ethical, theoretical, and methodological issues from other disciplines converge, KO is enriched and becomes more complex as these other disciplines encompass technology, linguistics, terminology, semiotics, and cultural and social studies. The internal discussion space within KO must address these diverse perspectives and influences, integrating them in a way that contributes to fulfilling its own objective: organizing human knowledge and making it available to everyone when they need to address their information needs.

Categorization and classification are central and interdependent processes responsible for building conceptual structures that support the representation and retrieval of information in Knowledge Organization Systems (KOS). In this context, it is not just about organizing

documents, but about stabilizing concepts, relationships, and points of view about the world, in order to make it intellectually navigable for specific communities of users (Hjørland 2016). Categorization and classification constitute a useful approach to describe a field as complex as KO, based on the assessment of its scientific production, whether at the local level or in relation to a specific country as is the case here.

2.1 Categorization in KO

Categorization is understood, first and foremost, as a cognitive process by which subjects group objects, phenomena, or documents based on perceptions of similarity and difference, often in a more flexible and situated way than formal classification. Jacob (2004) highlights that categorization systems establish order through non-binding associations between entities, constructed from the recognition of similarities in specific contexts, which gives them porous boundaries and the capacity to adapt to different uses and perspectives.

In KO, this cognitive dimension is indispensable: before placing a document in a structured system, the professional needs to recognize which set of meanings the informational object seems to belong to, based on categories such as theme, disciplinary domain, type of approach, among others (Dominguez Santana et al. 2024). In digital environments, marked by folksonomies, taxonomies, and flexible ontologies, categorization also becomes relevant, as it allows for the incorporation of multiple viewpoints, collaborative participation, and the dynamism of concepts, without the rigidity of a single, stable scheme.

2.2 Classification in the KO

Classification, in turn, is understood as a formalized process of distributing entities into classes according to explicit criteria and organized structures (hierarchies, facets, notations), as well as the system resulting from this process (classes, relations, and rules of use). Hjørland (2024) points out that KO, in a restricted sense, is especially dedicated to the development of KOS, which include classification schemes, subject heading lists, thesauri, and ontologies designed to describe and organize documents, subjects, and concepts.

In the field of Library Science, bibliographic classification organizes the universe of knowledge into areas and sub-areas in order to assign each document a defined position, favoring both systematic ordering and thematic retrieval. These systems, such as the Dewey Decimal Classification (DDC) or the Library of Congress Classification (LCC), employ controlled vocabulary, notations, and hierarchical relationships, frequently assuming the dual role of conceptually ordering a domain and physically (or logically) organizing collections. In many cases, such classifications reflect specific disciplinary and epistemological traditions, which demonstrates that no class structure is neutral; according to Hjørland (2002), the organization of knowledge is shaped by particular

interpretations of the relationships within domains, reflecting the needs and perspectives of specific communities.

2.3 Distinctions and complementarities between categorization and classification

Although often treated as synonyms in common language, categorization and classification can be conceptually distinguished. Jacob (2004) argues that classification, strictly speaking, tends to require mutually exclusive and exhaustive classes, defined by previously established principles, guaranteeing reference stability and predictability in the distribution of items. Categorization, on the other hand, operates more creatively and flexibly, allowing associations based on contextual similarity, in which the members of a category do not need to obey a single set of fixed criteria.

Hjørland (2002), however, observes that, in the current usage of the KO literature, classification often functions as an umbrella term that can encompass both the broad process of distinguishing and grouping entities and its more restricted and normative versions. From this perspective, categorization can be seen as the cognitive and linguistic plane of organization, while classification corresponds to the institutionalization of certain categorical choices in formalized KOS, with rules, notations, and application policies. Thus, both processes are complementary: there is no meaningful bibliographic classification without prior categorization, and categorization gains stability and power of social mediation when incorporated into classification schemes recognized by communities of practice.

2.4 The “classification” of the Knowledge Organization itself (IEKO/EIKO)

The field of KO itself has been the subject of efforts at epistemological categorization and classification. An Encyclopedia of Knowledge (IEKO), edited by Hjørland and Gnoli, seeks to map traditions, approaches, and paradigms in KO, describing, among others, practical intuitionist approaches, analytical facets, consensual approaches, cognitive user approaches, analytical/epistemological domain approaches. By identifying these "schools" and situating them within a conceptual framework, IEKO produces a kind of internal theoretical classification of KO, in which each approach represents a specific way of understanding what knowledge is, how it is organized, and how it should be represented.

The analytical domain approach, for example, argues that any object, document, or domain can be classified from multiple equally correct perspectives, according to the interests and paradigms of the communities involved (Mai 2011; Hjørland 2002). By gathering and describing such traditions, the IEKO demonstrates that KO can be organized according to epistemological categories (pragmatic, positivist, and historicist), showing that theories, paradigms, and approaches themselves function as KO systems. Ultimately, by "classifying" KO, these mappings reveal that all practices of categorization and classification—including those that structure the discipline itself—are traversed by theoretical commitments, social interests, and interpretive disputes.

3.0 Method

This research adopts a mixed-methods approach, combining qualitative and quantitative components, as it blends the interpretative analysis of content and thematic categories with the measurement of bibliometric indicators of scientific production in Knowledge Organization (KO). It is a first exploratory study, aimed at obtaining partial and descriptive results about the field, which will serve as a basis for subsequent investigations with complementary focuses and greater analytical refinement. The initial corpus consisted of works on KO indexed in BRAPCI; a total of 1,000 documents were retrieved, corresponding to the total number of records obtained with the search strategy: "knowledge organization" or "organization of knowledge," on February 15, 2026. The time frame covers the period from 1993 to 2025, in order to encompass all production indexed in BRAPCI and the context of the consolidation of KO as a structuring subfield of Brazilian Information Science.

In operational terms, the methodological procedure comprises: a) determination of the study universe, with BRAPCI being the chosen database for retrieving the documents to be analyzed; b) compilation and standardization of records of works related to KO, with removal of duplicates (44 duplicate documents were removed); and c) categorization of each work according to the thematic scheme of the ISKO Encyclopedia of Knowledge. Organization (IEKO),[1] which organizes the literature into six axes:

Table 1 EIKO Categories.

Code	Thematic Axis	Description (adapted from IEKO)
IEKO-A	Knowledge organization: general and historical issues	It covers the history of KO, its epistemological foundations, general theoretical perspectives, and the evolution of the field as a discipline.
IEKO-B	Central concepts in KO	It focuses on fundamental concepts such as classification, taxonomy, ontology, metadata, concept theory, and categorization.
IEKO-C	Knowledge organization systems (KOS)	It encompasses thesauri, classifications, ontologies, taxonomies, folksonomies, and other structured systems for organizing knowledge.
IEKO-D	Knowledge organization processes (KOP)	It deals with operational and methodological activities: indexing, cataloging, document classification, content analysis, and descriptive representation.
IEKO-E	Methods, approaches, and philosophies	It brings together theoretical perspectives, methodological approaches (e.g., domain analysis, concept theory), and philosophical reflections on organizational communication.
IEKO-F	KO in different contexts and applications	It applies the principles and systems of KO to specific domains: business, education, legal, health, digital, among others.

Source: Prepared by the authors, based on the EIKO category, 2026.

The adoption of the IEKO categorization proves particularly suitable as it integrates principles from Hjørland (2016)—especially the distinction between organizational

1 https://www.isko.org/cyclo/ .

processes and systems—and the *Classification System for Knowledge Organization Literature* (CSKOL), in addition to allowing the results obtained to be directly comparable with international mappings of the field.

3.1 Thematic Categorization Method

The classification of the documents followed the thematic scheme of the *ISKO Encyclopedia of Knowledge Organization* (IEKO), composed of six axes: (a) general and historical issues; (b) central concepts; (c) knowledge organization systems; (d) knowledge organization processes; (e) methods, approaches, and philosophies; and (f) KO in different contexts and applications. Six complementary categories were added—Information technologies and systems, Organizational knowledge management, Metric studies and scientific production, Information policies and social aspects, Teaching and training in KO and Undetermined not applicable and teaching in KO – to encompass recurring themes in Brazilian production.

The thematic classification of the documents was carried out based on an automated procedure supported by terminological dictionaries constructed for each of the axes of the International Society for Knowledge Organization (ISKO). These dictionaries were developed from a review of the specialized literature on Knowledge Organization, an exploratory analysis of the title, keyword, and abstract fields of the analyzed documents, as well as consultation of terminological instruments in the field, such as the BRAPCI Thesaurus of Information Science. For each thematic axis, sets of representative terms were defined in Portuguese, Spanish, and English, allowing for the automatic identification of correspondences in bibliographic records through searches insensitive to differences in accentuation and capitalization.

The assignment of thematic categories considered a weighting system of the terms identified in the different fields of the records. In order to prioritize information with greater thematic representativeness, differentiated weights were assigned to the occurrences found in each field: keywords (weight 3), title (weight 2), and abstract (weight 1). This strategy is based on the understanding that keywords represent the concepts explicitly selected by the authors to characterize the document's content, while titles synthesize its main focus and abstracts provide a broader contextualization, subject to greater terminological diversity. The definition of these weights was initially based on classic studies of automatic indexing and textual classification (Salton and Buckley 1988; Sebastiani 2002) and subsequently validated through a pre-test with 50 documents. Different weighting strategies were compared, and the one that showed the greatest agreement with the manual reference classification was adopted (Kappa = 0.82). The final category of each document was assigned based on the highest accumulated score among the analyzed axes.

As a complementary validation mechanism, confidence levels were established for the classifications obtained. Confidence was determined by the difference between the scores of the two highest-ranked categories in each document: high differences indicated high

confidence, intermediate differences indicated medium confidence, and small differences were classified as low confidence. Records falling under this last condition received priority attention in the manual verification stage. To assess the consistency of the procedure, a stratified sample corresponding to 10% of the corpus (95 documents) was selected, encompassing proportional representation of thematic categories, different publication periods, and distinct document types. The documents were independently reclassified by two researchers specializing in Knowledge Organization, without prior access to the results of the automatic classification. The agreement between the coders was considered high (Kappa = 0.86), indicating almost perfect consistency according to the criteria of Landis and Koch (1977). Furthermore, a 79.3% agreement was observed between the automatic classification and the researchers' evaluation, reinforcing the reliability of the method. Divergent cases were subsequently discussed until consensus was reached, allowing for successive refinements in the terminological dictionary and contributing to increased robustness and transparency of the classification process.

Automated term analysis was conducted using data processed in Google Sheets and operationalized through an AppScript developed with DeepSeek v3.2 (free version), combined with a controlled categorization approach.

The categorization process comprised three stages: (a) conceptual extraction, based on titles, keywords, and abstracts; (b) semantic mapping, involving the alignment of extracted terms with ISKO categories through terminological correspondence and conceptual proximity; and (c) final classification, consisting of the assignment of a primary category and a representative subcategory. Data were subsequently structured and organized in an Excel spreadsheet. Calculation of absolute and relative frequency and comparative analysis between categories. In accordance with the principles of open science and replicability advocated by the ISKO community, the spreadsheet with all classified records and the scripts used are publicly available at the following address: [https://docs.google.com/spreadsheets/d/1ptI6xE2NLi4XA-xcu59rxcw3BshffmdSP1NnvzHr_cM/]. The source code of the script in Google Apps Script can be accessed and reused upon request to the authors.

In this context, the use of a bibliometric study applied to the field of KO is justified, since this type of approach offers rigorous instruments to describe, in quantitative terms, the temporal evolution, the channels of scientific communication, as well as the thematic distribution of production. By articulating bibliometric indicators with the specialized categorization of IEKO, it is possible not only to map the internal configuration of KO in Brazil, but also to identify concentrations and gaps in certain sub-areas, asymmetries in the participation of authors and institutions, and potential fronts for theoretical and applied expansion. In a domain whose historical and conceptual development is in the process of consolidation, bibliometrics therefore contributes to highlighting patterns underlying scientific activity and to situating Brazilian production in dialogue with international literature, providing subsidies for strategic research planning and for strengthening the organization of knowledge as a disciplinary field.

Limitations of the method include its reliance on metadata (not full text), potential conceptual ambiguity, and predominantly semi-automatic classification. Corpus specificity: the results reflect the production indexed in BRAPCI and, therefore, are representative of the Brazilian context, not necessarily generalizable to other regions or databases. It is recommended that the results be interpreted as a systematic approximation of the thematic distribution of Brazilian scientific production in KO, complemented, when pertinent, by in-depth qualitative analyses.

4.0 Results

The scientific production retrieved from BRAPCI presented 956 works distributed over the years, with the first publications dating back to 1993. These articles were published in the journal Ciência da Informação, authored by Clovis Ricardo Montenegro de Lima with an article entitled: *AIDS - the epidemics of viruses and information, and an article by Clarinda Rodrigues Lucas, entitled "The organization of knowledge and information technologies". This marked the beginning of studies in KO (Organization of Knowledge)*.

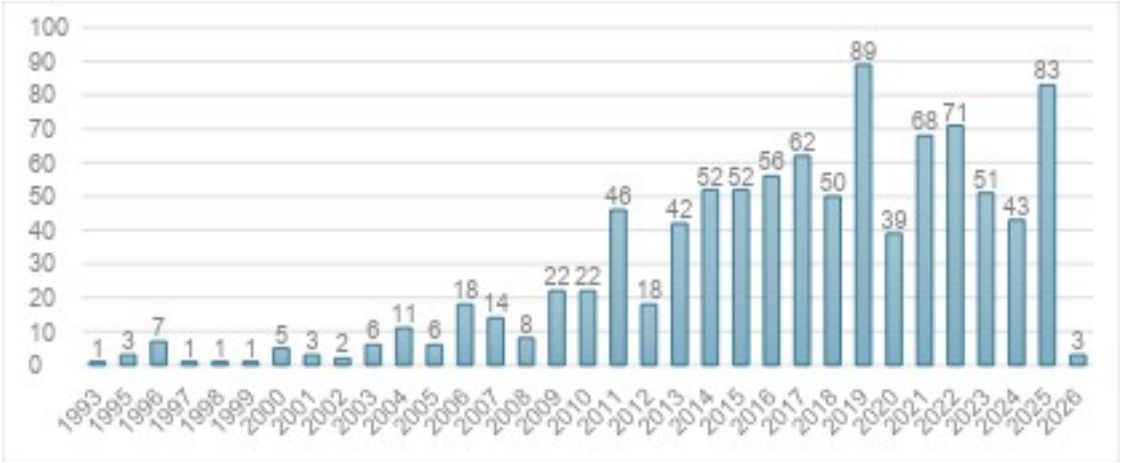

Figure 1. Graph showing the evolution of scientific production in KO.

A greater growth in productions using a KO approach is observed from 2004 to 2025, with the most productive years being from 2011 to 2025.

The analysis of the distribution of the 956 classified documents revealed a thematic concentration in three main areas, which together account for 58.5% of the total production: Methods, approaches and philosophies with 20.9%, Central concepts in KO with 20.6% and KO in different contexts and applications with 17.0%.

This pattern highlights a Brazilian scientific output that simultaneously prioritizes:

- Theoretical-methodological reflection

- The conceptual foundation
- Application in diverse contexts

On the other hand, the axes focused on more operational and instrumental aspects—Processes and Systems—presented relatively low percentages (6.5% and 6.0%, respectively), suggesting less emphasis on studies of the practice of KO and on the documentary artifacts themselves.

Table 2 Distribution of retrieved documents by EIKO Category.

EIKO Category	**Absolute Frequency (n)**	**Relative Frequency (%)**
Methods, approaches and philosophies	200	20.92%
Core concepts in KO	197	20.60%
KO in different contexts and applications	162	16.95%
KO: General and historical issues	67	7.01%
Knowledge organization processes (KOP)	62	6.49%
Knowledge organization systems (KOS)	57	5.96%
Other suggested categories	211	22.07%

Source: research data extracted from BRAPCI, 2026.

The residual category “Other suggested categories” (22.07%) proved to be heterogeneous and epistemologically relevant. Upon examining these 211 documents, the interpretative analysis allowed for the identification of at least six emerging sub-themes:

Table 3 Distribution of retrieved documents by thematic grouping, suggesting new categories.

Other suggested categories	**Absolute Frequency (n)**	**Relative Frequency (%)**
Information technologies and systems	77	8.05%
Undetermined / Not applicable	52	5.44%
Information policies and social aspects	30	3.14%
Metric studies and scientific production	28	2.93%
Teaching and training in KO	18	1.88%
Knowledge Organizational management	6	0.63%

Source: research data extracted from BRAPCI, 2026.

The analysis of the category "Other suggested categories" reveals a process of reconfiguration in Knowledge Organization, marked by the emergence of new thematic axes, and by the growing attention to the interdisciplinary aspects of KO. The most prominent are Information Technologies and Systems (8.05%) and Information Policies and Social Aspects (3.14%), followed by Metric Studies and Scientific Production (2.93%) and Education and Training in KO (1.88%). The sum of these latter categories presents a percentage equivalent to that of the category "Undetermined / Not applicable" (5.44%), suggesting classification limitations or the existence of hybrid and interdisciplinary research outputs.

5.0 Discussion

The results indicate a predominantly theoretical-conceptual approach to Knowledge Organization (KO) in Brazil, with an emphasis on epistemological foundations and methodological reflection, to the detriment of research on artifacts and operational procedures in the field. This concentration on the axes of Methods, approaches and philosophies, Central concepts in KO and KO in different contexts aligns with international movements that broaden the scope of KO beyond its library science roots, incorporating critical, social and applied perspectives (Hjørland 2017).

One of the most relevant findings is the identification of emerging categories—such as technologies applied to KO, digital semantic infrastructures, social and critical dimensions of information, knowledge management in digital environments, and education and training in KO—which indicate a process of reconfiguration of the field, in which new research agendas begin to coexist with traditional axes. Of particular note is the significant output on the influence and application of technologies in our field.

In this context, the IEKO taxonomy, while effective for classic domains, shows limitations in representing these new dynamics. The KO category in different contexts and applications, for example, proved to be excessively broad, functioning as a residual grouping for applied studies without better framing. Detailed analysis allowed the identification of distinct subgroups – applications in libraries, digital environments, organizational contexts, and hybrid approaches – highlighting the need for greater classification granularity.

The data thus reveal a science in transition: Brazilian KO maintains solid theoretical foundations, expands into new contexts and technologies, but reduces its focus on its classic processes and systems. This movement can be interpreted as a disciplinary maturation, in which the area seeks legitimacy through theoretical sophistication and the expansion of its boundaries.

Overall, the results point to a field in which consolidated conceptual structures coexist with new demands imposed by contemporary technological, social, and epistemological transformations. This coexistence puts strain on traditional classification models and reinforces the need for their updating.

6.0 Conclusions

This study aimed to analyze the scientific production in the field of Knowledge Organization in light of the classic ISKO categorization, additionally incorporating an empirical approach based on content analysis to identify emerging themes. The results highlighted both the consolidated structure of the field and its dynamics of transformation, characterized by a hybrid profile: strongly anchored in theoretical and conceptual foundations, but also attentive to applications and technological innovations. The classic axes of KO (processes and systems) appear with reduced weight, while interfaces with information policies, metric studies and technologies gain relevance.

This profile suggests a field undergoing a process of reconfiguration, seeking to balance disciplinary tradition with the contemporary demands of the information society. The data invite reflection on how different national traditions of KO position themselves in relation to the structuring axes proposed by IEKO, revealing specificities that enrich the international panorama of the area.

As a contribution, this study proposes a broader reading of the area, grounded in empirical evidence, which allows not only a description of the field's structure but also the identification of its evolutionary trends. In this sense, it reinforces the importance of revisiting and updating traditional theoretical and classificatory models in order to keep pace with contemporary transformations in scientific production.

Finally, it is suggested that future research further analyze the identified emerging categories, as well as explore complementary methodologies—such as network analysis and semantic modeling—to broaden the understanding of the dynamics of the KO field.

Acknowledgements

We sincerely appreciate and acknowledge the suggestions made by the reviewer(s) and Editor for improvements and final shape of the paper. We thank CAPES—Coordination for the Improvement of Higher Education Personnel (Brazil)—for the funding (Finance Code 001) associated with Process No. 88887.340098/2026-00 of the PIPD Program.

Declaration on the use of artificial intelligence

It is hereby declared that, during the development of this work, the generative Artificial Intelligence tool DeepSeek was used exclusively as a technical support resource for the development of programming code employed in the automatic categorization of the data analyzed.

The use of the tool was limited to generating and assisting in the structuring of the computational code and was not employed in the development of the intellectual content of the work, interpretation of the results, formulation of the analyses and conclusions, or authorial writing. The definition of the categorization criteria, methodological application,

validation of the procedures, and interpretation of the results remained under the responsibility of the authors.

The authors assume full responsibility for the content of the work, including the methodological procedures, results, analyses, and conclusions presented.

References

Alves, B. H., and Ely Francica Tannuri Oliveira. 2016. “O desenvolvimento do domínio da "organização do conhecimento" no contexto da Ciência da Informação a partir da ISKO-Brasil.” *Brazilian Journal of Information Science* 10, no. 2: 103-108.

Dominguez Santana, Lílian, Raissa Martins, Leonardo Chagas, Frederico Pereira, Gercina Lima and Maria Aparecida Moura. 2024. “Sistemas de Organização do Conhecimento: Análise Comparativa e Modelagem de Instrumentos de representação do Conhecimento”. Encontros Bibli: Revista eletrônica de Biblioteconomia e Ciência da informação 29. Florianópolis/SC, Brasil: 1-27. https://doi.org/10.5007/1518-2924.2024.e97708.

Farías Blundell, Verónica, Natalia Aguirre-Ligüera and Mário Barité. 2022. “Aporte latinoamericano a la investigación en Organización del Conocimiento: una aproximación bibliométrica.” Bibliotecas. Anales de Investigación 18, no. 1: 109-129.

Guimarães, José Augusto Chaves. 2020. “Domains, Contexts and Foundations of Knowledge Organization.” Perspectives in Information Science 25, no. 3: 4–20.

Hjørland, Birger. 2002. “Domain Analysis in Information Science: Eleven Approaches—Traditional as Well as Innovative.” Journal of Documentation 58, no. 4: 422–62.

Hjørland, Birger. 2016. “Knowledge Organization (KO): What Is It About?” Knowledge Organization 43, no. 6: 475–484.

Hjørland, Birger. 2017. “Classification.” Knowledge Organization 44: 97–128. https://www.inlibra.com/document/download/pdf/uuid/988a152b-beef-31d7-bd4a-beb2b402c79d

Hjørland, Birger. 2024. “Knowledge Organization.” In ISKO Encyclopedia of Knowledge Organization, edited by Birger Hjørland and Claudio Gnoli. https://www.isko.org/cyclo/knowledge_organization.

Hjørland, Birger and Claudio Gnoli, eds. n.d. “ISKO Encyclopedia of Knowledge Organization”. https://www.isko.org/cyclo/.

Jacob, Elin K. 2004. “Classification and Categorization: A Difference That Makes a Difference.” Library Trends 52: 515–540. https://hdl.handle.net/2142/1686.

Landis, J. Richard, and Gary G. Koch. 1977. “The Measurement of Observer Agreement for Categorical Data.” Biometrics 33, no. 1: 159–174.

Mai, Jens-Erik. 2011. “The Modernity of Classification.” Journal of Documentation 67, no. 4: 710–30.

Salton, Gerard, and Christopher Buckley. 1988. “Term-Weighting Approaches in Automatic Text Retrieval.” Information Processing & Management 24, no. 5: 513–523.

Sebastiani, Fabrizio. 2002. “Machine Learning in Automated Text Categorization.” *ACM Computing Surveys* 34, no. 1: 1–47.

Souza, A. M. 2013. “Scientific Production in Knowledge Organization: A Bibliometric Analysis.” Master's thesis, Federal University of Minas Gerais, Belo Horizonte.

Villalobos, Ana Paula, and Levi Alã Neves dos Santos. 2015. “Análise da produção científica brasileira: Uma contribuição da mineração de dados para a Ciência da Informação”. Tendências da Pesquisa Brasileira em Ciência da Informação 8, no. 1: 34-48. https://revistas.ancib.org/tpbci/article/view/367.