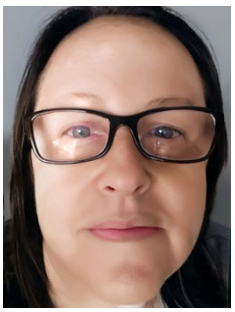
# Knowledge Organization Systems in Reference Works Specialized in Knowledge Organization: A Study of Punctual Terminology

Mario Barité*, Varenka Parentelli**, and Mirtha Rauch***

* ** *** University of the Republic, Uruguay. Address: San Salvador 1944, CP 11200, Montevideo, Uruguay
*mario.barite@fic.edu.uy, ** varenka.parentelli@fic.edu.uy, *** mbrauch@gmail.com

Mario Barité is Full Professor at the Faculty of Information and Communication, University of the Republic, Uruguay. He holds a PhD and has a Master's degree in Scientific Information from the University of Granada, Spain. His main research interests include knowledge organization and terminology. He is a researcher of the Uruguayan National System of Researchers. Among another contributions, he has published a *Dictionary of Knowledge Organization*, the sixth edition of which is available in Internet. He is editor-in-chief of *Informatio* journal.

Varenka Parentelli is Professor and Academic Coordinator of the APEX Program and Professor of the Faculty of Information and Communication, University of the Republic, Uruguay. PhD in Education Sciences (National University of La Plata), Master's Degree in University Teaching and Bachelor's Degree in Communication Sciences (University of the Republic), Diploma in Educational Management and File Management, Research and Pedagogical Uses (FLACSO, Argentina). His main lines of research include higher education, interdisciplinarity and image. She is a researcher of the National System of Researchers of Uruguay. He has made numerous publications. He is part of the Board of the magazine *IT Salud Comunitaria y Sociedad*.



Mirtha Rauch is a sworn translator for French and English into Spanish and vice versa. At present, she works as Assistant Professor for Professional Practice at the Public Translation Department, School of Law, University of the Republic of Uruguay. She also is a librarian, specialized in the health area, with major experience in neurology and neurosurgery. She has been involved in terminology research for the last twenty years and is now part of the group devoted to knowledge organization and terminology in the School of Information of the University of the Republic, Uruguay.

Barité, Mario, Varenka Parentelli, and Mirtha Rauch. 2024. "Knowledge Organization Systems in Reference Works Specialized in Knowledge Organization: A Study Of Punctual Terminology". *Knowledge Organization* 51, no. 5: 355-361. 29 references. DOI:10.5771/0943-7444-2024-5-355.

**Abstract:** Specialized reference works operate as catalysts and organizers of the terminological processes involved in the evolution of specialized concepts. This research aims to identify the term 'knowledge organization system' records and its synonymous or similar expressions in dictionaries, glossaries, vocabularies, and encyclopedias specialized in KO published over time. The research is qualitative and constitutes a study of punctual terminology, which introduces the perspective of dictionarization studies. The results yield fourteen terms representing more or less the same concept registered in the twelve reference works found. Among other conclusions, it is assumed that there has been progress towards terminological unification around the expression 'knowledge organization system'.

**Received**: 15 June 2024; Accepted 13 July 2024.

**Keywords**: knowledge organization system; terminology; reference works.

† This article was selected as one of the best papers at the Eighteenth International ISKO Conference, March 20-22, 2024, in Wuhan, China.

## 1.0 Introduction

Knowledge organization system is a generic term that combines a set of vocabularies and classifications, universal or specialized in scope. As Mazzocchi (2018, 54) points out, "they are characterized by different specific structures and functions, varied ways of relating to technology, and used in a plurality of contexts by diverse communities", but they all have in common "that they have been designed to support the organization of knowledge and information to make their management and retrieval easier" (Mazzocchi 2018, 54).

In the English literature on knowledge organization (KO) until the end of the twentieth century, various expressions (such as indexing language, documentary language, controlled vocabulary, structured vocabulary, and classification system, among others) competed to assume the semantic representation of the set of systems or controlled languages, created to sustain conceptual structures for general classification and/or indexing of documents and information.

It was not until 1998 that the term 'knowledge organization system' (henceforth, KOS) appeared in the title of a Conference held in Pittsburgh, Pennsylvania: 'Networked Knowledge Organization Systems (NKOS)'. In 2000, Hodge used it and provided the first systematization of its concept. However, San Segundo (1996) had already used the term in Spanish (*sistema de organización del conocimiento*), although restricting its scope to classification systems.

Contemporary terminology that seeks to specify or provide new approaches to studying KOS begins to appear. For example, Zang and Mayr (2019, 3) use the expression LOD KOS "as an umbrella term to refer to all value vocabularies and lightweight ontologies within the Semantic Web framework".

There is no unanimous opinion regarding the types of systems that can be considered KOS. Hodge (2000) established three KOS categories, each one with its specific types: lists (authority files, glossaries, dictionaries, gazetteers); classifications and categories (subject headings, classification schemes, taxonomies, categorization schemes); and relationship lists (thesauri, semantic networks, and ontologies).

The North American standard (National Information Standards Organization 2005) only mentions four types of KOS: lists, synonym rings, taxonomies, and thesauri. In the keywords that he introduces to his work, Soergel (2009) lists the following: ontologies, metadata schema, taxonomies, classifications, web directory structures, filing plans, thesaurus / thesauri, dictionaries, folksonomies, authority files, gazetteers, faceted classifications, and subject headings. Later, when classifying the KOS by generic function, by content and structure, and by origin and editorial control, this author adds to the previous enumeration authority lists, coding schemes, lists of synonym rings, glossaries, concept maps, mind maps, semantic networks, RDF graphs, and topic maps (Soergel 2009, 5-6). Finally, Abbas (2010, 99-131) distinguishes between metadata structures and content creation standards (machine-readable cataloging and other metadata schemas), controlled vocabularies (subject headings lists and thesauri), and classification schemes (Dewey Decimal Classification, Library of Congress Classification, Colon Classification Scheme and others).

Beyond these divergences, the common point is that all KOS

> [..] are used by people to find information and make sense of it; KOS must support people in their quest for meaning; they must present meaningful structures of concepts. KOS are also used by computer programs to reason about data; KOS must represent formal knowledge about concepts (Soergel 2009, 3).

Specialized reference works (dictionaries, glossaries, encyclopedias, vocabularies of different types and dimensions) operate as catalysts and organizers of the terminological processes involved in evolving specialized concepts. In these reference works, the dictionarization processes of each term are recorded, as well as the different conceptions that terminologists capture in the graphic form of the terms, in their choice of preferred terms, and their definitions.

The term 'dictionarization' was coined by Nunes with the scope of "description and instrumentalization of the language based on the dictionary" (Nunes 2002, 99). Nunes himself further specified the concept by pointing out that "the study of dictionarization implies making explicit the historical processes that lead to the formation of [the dictionaries], as well as showing the appearance and transformations of the practices that allow its construction" (Nunes 2006, 45).

The specificity of dictionarization studies appears in the exclusive use of dictionaries and other reference works as an object of terminological study. They are based on the fact that the reference works have a specific date of publication, which makes it possible to analyze the historical record of the terminology.

Dictionary studies help to identify the traceability of terms and their concepts. Traceability is understood as the process that begins with the coining and continues in the study of use, documentation, and the eventual modification or extinction of the terms.

The choice of terms to be incorporated into a specialized reference work may be based on various criteria, the consistent application of which favors the internal coherence of the work: i) the criterion of use, which favors the selection of terms actually used by specialists and other members of a discourse community; ii) the criterion of specificity, which focuses only on terms that are seen as proper and exclusive to the domain; and, iii) the documentation criterion, that is,

the appearance of specialized terms in prototypical documentary types of the area, such as journals, proceedings, manuals, educational and popularization texts, as well as other reference works.

The three criteria mentioned (usage, specificity, documentation) largely correspond to user warrant (Lancaster 1977), academic warrant (Svenonius 2003), and literary warrant (Beghtol 1986), respectively. The consistent use of one warrant or another to include or exclude the terms registered in specialized dictionaries contributes to reasonable and balanced thematic coverage.

Even so, any comparison between dictionaries of the same domain will reveal differences in length, depth, the number of terms chosen, or how they formalize their terminology. This should be interpreted as a normal situation. Lukasik (2017, 4-5) identifies a dozen differentiation factors in the content of reference works in the same domain, among which stand out: "culture-dependent terminology", regional variants, "various degrees of formality and informational density", "different styles of reasoning", "the existence of idiosyncratic terminological/conceptual systems", "some new terms entering the lexicon, and — most importantly — some terms changing their definition, scope or applicability."

In terminology, studies of punctual terminology -focusing on various linguistic, grammatical, terminological, and translation aspects of one term or a small number of related terms- are usual. The studies of punctual terminology collaborate methodologically in the dictionarization analyses. In this work, a study of specific terminology will be articulated with a study of dictionarization, focused on the term 'knowledge organization system' and its related terms.

## 2.0 Objectives

This research aims to identify the scope and exhaustiveness with which the term 'knowledge organization system', its acronym (KOS), and its synonymous or similar expressions have been registered in dictionaries, glossaries, vocabularies, and specialized encyclopedias in published KOS over time. This objective is intended to determine the traceability of the generic terms used in the area to refer to the systems with greater or lesser vocabulary control.

The following specific objectives are also established: a) promote the development of terminological studies in KO, not only from the perspective of domain analysis but also, as in this case, from punctual terminology approaches; b) bring dictionarization studies closer to KO, considering the interdisciplinary nature of both fields; c) contribute to the study of the evolution of the concepts of KO.

## 3.0 Methodology

The approach carried out for this work is qualitative and constitutes a study of punctual terminology. Punctual terminological research consists of "a technique for investigating a term or a restricted group of terms belonging to any field of human activity, with a view to satisfying an immediate need expressed by a user" (Célestin et al. 1984, 17). As mentioned by Barité and Rauch (2022, 6),

> [...] these authors also establish a typology of punctual research, which includes the analysis of the meaning of a term, the most appropriate term to represent a new concept, the certification or validation of the use of a term, or the search for an exact or approximate equivalence in another language.

Generally, punctual terminology studies are used to solve problems of designation, conceptualization, translation, and historical or etymological precedence, continuously presented by neologisms. The demands and urgencies posed by the avalanche of neologisms to translators, communicators, journalists, and information professionals make punctual research helpful for selecting terms in target languages or descriptors or indexing terms for retrieving information in native languages.

This work, as said, also introduces the perspective of dictionarization studies to the extent that the research is carried out using specialized dictionaries, encyclopedias, and glossaries as the object of study.

For this work, the following methodological phases were fulfilled:

i) *Formation of the corpus*. It was integrated with all specialized KO dictionaries, encyclopedias, and glossaries published to date in various parts of the world and different languages. Reference works partially overlapping with other areas, such as organization and information processing, were included as long as the thematic predominance is KO. General reference works from Information Science, Library Science, or Documentation and specialized reference works only in Classification (former name of the KO domain) were excluded. Works specialized in KO subdisciplines, such as indexing, or related areas, such as information retrieval, were also excluded, especially since these do not include terms on KOS but on processes and information retrieval. From the search carried out in different databases, twelve (12) specialized works on KO published between 1997 and 2023 were identified and distributed as follows: an encyclopedia (Hjørland and Gnoli 2017), two dictionaries (Satija 2004; Barité et al. 2015), and nine glossaries. Of the latter, four glossaries are autonomous (Barité 1997; Wellisch 2000; Gnoli et al. 2006; Bonotto 2007), three are glossaries attached to standards (NISO 2005; BSI 2005-

2008; ISO 2011-2013), and two are classification system annexes (Bliss 2016; OCLC 2023).

*ii) Compilation of terms*. All dictionaries and glossaries have a main alphabetical organization, while the encyclopedia has a systematic organization and an auxiliary alphabetical index. The technique used was to go through the alphabetic structures from beginning to end, compiling generic terms that responded to the contemporary notion of a knowledge organization system. The definitions were read in all cases to ensure the meanings correspond to the concept studied. Preferred terms (identified by having a definition) and non-preferred terms (presented without definition and with a reference to the preferred term as input) were collected.

iii*) Indication of the terms registered by each source*.

iv) *Distribution of the terms according to the number of times they appear in the sources and the years of publication of these sources to establish the traceability of said terms*.

## 4.0 Results

Table 1 orders the sources chronologically and indicates the terms found in each. Expressions from works published in Spanish, Portuguese, and Italian were translated into English, preferably using the equivalence provided in the same works. Terms not preferred by the authors of the reference works are indicated in parentheses.

As shown in Table 1, fourteen (14) generic terms reasonably associated or equivalent to the current concept of knowledge organization system were found. The diversity of denominations for the same concept or very close concepts is significantly high. In one of the sources, up to eight generic terms are mentioned, and in another, only one.

In all the reference works, terms relevant to the research were found, which should not be surprising, considering that according to Hjørland (2008) the conceptual structure of KO is organized around two main axes, one of which is knowledge organization systems.

Regarding the particular analysis of some terms, 'classification scheme' and 'schedule' (whose conceptual scope is not unanimously accepted) were included because at least one of the authors assigns a broader generality than the reference to classification systems.

The terms 'controlled vocabulary' and 'indexing language' appear mentioned throughout the entire period covered by the works. If the research were to be expanded, they might appear mentioned in dictionaries and other reference works on Information Science since the 1960s onwards.

For its part, 'documentary language' is a more or less literal translation of the French '*langage documentaire*', surely derived from Documentation, as a name for Library Science studies in France. The first record of the term in English reference works that has been found is almost fifty years old (Wersig and Neveling 1976, 67). However, it seems to have

| Sources | Generic terms for KOS |
|---|---|
| Barité (1997) | Controlled vocabulary. Documentary language. Indexing language. Indexing system. |
| Wellisch (2000) | Classification scheme. Controlled vocabulary. Indexing language. Schedule. |
| Satija (2004) | Classification scheme. Controlled vocabulary. Documentary language. Index language. (Indexing language). Schedule. |
| BSI (2005-2008) | Classification scheme. Controlled vocabulary. Schedule. Structured vocabulary. |
| NISO (2005) | Classification scheme. Controlled vocabulary. Indexing language. |
| Gnoli *et al*. (2006) | Classification scheme. Controlled vocabulary. Indexing language. Knowledge organization system. (KOS). Schedule. |
| Bonotto (2007) | (Classification scheme). Controlled language. Controlled vocabulary. Documentary language. (Indexing language). |
| ISO (2011-2013) | Controlled vocabulary. Structured vocabulary |
| Barité *et al*. (2015) | Controlled vocabulary. (Controlled language). Documentary language. Indexing language. Knowledge organization system. (Knowledge organization and representation system). (KOS). (KORS). |
| Bliss (2016) | Classification scheme. Controlled index language. Schedule. |
| Hjørland and Gnoli, (2017-...) | Knowledge organization systems (KOS). |
| OCLC (2023) | Schedule. |

*Table 1.* Terms in KO reference works.

been used more in countries with a Latin culture (France, Spain, Latin America) for its equivalents in French, Spanish, and Portuguese, as arises from the regional origin of the reference works studied.

The term 'uncontrolled vocabulary' appeared in a work (Wellisch 2000), but it was not considered because the most widespread idea is that KOS have, to a greater or lesser extent, forms of vocabulary control (although this may be debatable).

Table 2 presents the fourteen (14) terms compiled in alphabetical order, indicating the number of entries in the dictionaries and the year of registration in the sources. The years that appear in parentheses correspond to the registration of each expression as a non-preferred term.

A predominance of the term 'controlled vocabulary' is clearly visualized throughout the entire period, with 9 records out of a total of 12 works (75%). Classification scheme and Indexing language, for their part, appear in 7 works (58.3%) and Schedule in 6 (50%). On the other hand, two acronyms were identified (KOS, KORS), which in their four records are considered non-preferred terms.

It is worth mentioning that 6 of the 14 terms only appear once, which could indicate little recognition within the domain or the existence of relatively new terms that have not yet been consolidated in the area.

The set of results could be contrasted with studies of similar specific terminology that, based on these lists, take into account other segments of the specialized literature on KO, such as journal articles, communications in proceedings, books, or manuals.

## 5.0 Conclusions

The terminological construction, modification, and deconstruction processes entail the sudden incorporation of new terms and the obsolescence of old ones, especially in areas that are developing with certain dynamism, such as KO. In the flow of these processes, the authors try to establish consensus, mark their conceptual differences, and even propose new terminology that better reflects the essence of their ideas.

When several specialized dictionaries and glossaries are available in a domain, it is quickly discovered that they are all different in terms of the terminology they select, record, and explain. Implicitly, this means that the authors have constructed different interpretations of the domain's terminological structure.

The criteria chosen to select terms are important in administering these differences. However, the decisions made regarding the domain map also impact the notional structure configured from the relationships between the selected terms.

In the case of KO's reference works, still few, perhaps due to the relative youth of the field, the study of expressions more or less equivalent to the contemporary 'knowledge organization system' shows a large (and even excessive) deploy-

| Terms | Entries | Years of sources |
|---|---|---|
| Controlled vocabulary | 9 | 1997, 2000, 2004, 2005, 2005-2008, 2006, 2007, 2011-2013, 2015 |
| Classification scheme | 7 | 2000, 2004, 2005, 2005-2008, 2006, (2007), 2016 |
| Indexing language | 7 | 1997, 2000, (2004), 2005, 2006, (2007), 2015 |
| Schedule | 6 | 2000, (2004), 2005, 2006, 2016, 2017 |
| Documentary language | 4 | 1997, 2004, 2007, 2015 |
| Knowledge organization system | 3 | 2006, 2015, 2017 |
| KOS | 3 | (2006), (2007), (2017) |
| Structured vocabulary | 2 | 2005-2008, 2011-2013 |
| Controlled index language | 1 | 2016 |
| Index language | 1 | 2004 |
| Indexing system | 1 | 1997 |
| Controlled language | 1 | 2007 |
| Knowledge organization & representation system | 1 | (2015) |
| KORS | 1 | (2015) |

*Table 2.* Generic Terms for KOS.

ment of terms to represent more or less the same idea, as well as divergences in their conceptual delimitation.

These difficulties herald that some of these equivalent expressions will continue to be used interchangeably for some time because they carry their own history behind them. Meanwhile, the process towards terminological unification around the expression 'knowledge organization system' and its literal translations into other languages may advance.

**Acknowledgment**

This work was carried out with the support of the Sectoral Scientific Research Commission of the University of the Republic of Uruguay.

**References**

Abbas, June. 2010. *Structures for Organizing Knowledge Exploring Taxonomies, Ontologies, and Other Schemas.* Chicago: Neal-Schuman Publishers.

Barité, Mario. 1997. *Glosario sobre Organización y Representación del Conocimiento: Clasificación, Indización, Terminología*. Montevideo: CSIC-EUBCA.

Barité, Mario, Stephanie Colombo, Amanda Duarte Blanco, Lucía Simón, Gabriela Cabrera Castromán, María Luisa Odella, and Mario Vergara. 2015. *Diccionario de Organización del Conocimiento: Clasificación, Indización, Terminología*. 6ª ed. corregida y aumentada. Montevideo: CSIC.

Barité, Mario, and Mirtha Rauch. 2022. "Terminological Studies as Domain Analysis: A Critical Exploration." *Brazilian Journal of Information Science* 16, e02140. https://revistas.marilia.unesp.br/index.php/bjis/article/view/12601/8644

Beghtol, Clare. 1986. "Semantic validity: Concepts of Warrant in Bibliographic Classification Systems." *Library Resources & Technical Services* 30, no. 2: 109-123.

Bliss Classification. 2016. "Glossary". In *Bliss Classification*. *Overall Introduction to the Scheme.* https://www.blissclassification.org.uk/bcsched.shtml

Bonotto, Marta E.K. Kling. 2007. *Glossário da Área de Organização e Tratamento da Informação.* UFRGS, [20--].

British Standards Institution. 2005-2008. *Structured Vocabularies for Information Retrieval: BS 8723*-.S.l.: BSI.

Célestin, Tina, Gilles Godbout, and Pierrette Vachon-L'Hereux. 1984. *Méthodologie de la Recherche Terminologique Ponctuelle*. Québec: Office de la Langue Française.

Gnoli, Claudio, Vittorio Marino, and Luca Rosati. 2006 *Glossario di Organizzazione della Conoscenza*. http://www.iskoi.org/doc/glossario.htm

Hodge, Gail. 2000. *Systems of Knowledge Organization for Digital Libraries: Beyond Traditional Authority files*. Washington D.C.: Council on Library and Information Resources. http://www.clir.org/pubs/reports/pub91/contents.html

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

International Organization for Standardization. 2011-2013. *Information and Documentation : Thesauri and Interoperability with Other Vocabularies: ISO 25964-1: ISO 25964-2.* Vernier: ISO.

Hjørland, Birger, and Claudio Gnoli, eds. 2017. *ISKO Encyclopedia of Knowledge Organization (EIKO).* https://www.isko.org/cyclo/

Lancaster, Frederick Wilfrid. 1977. "Vocabulary Control in Information Retrieval Systems." In *Advances in Librarianship*, 1-40, edited by Melvin Voight and Michael Harris. London: Academic Press. v. 7.

Łukasik, Marek. 2017. "Contrastive Terminography". *Cognitive Studies / Études cognitives*, v. 17, 1-14. https://ispan.waw.pl/journals/index.php/cs-ec/article/view/cs.1378/3060

Mazzocchi, Fulvio. 2018. "Knowledge Organization System (KOS): An Introductory Critical Account." *Knowledge Organization* 45, no. 1: 54-78. doi.org/10.5771/0943-7444-2018-1-54

National Information Standards Organization. 2005. Guidelines for the Construction,

Format, and Management of Monolingual Controlled Vocabularies. ANSI/NISO Z39.19-2005. Bethesda: NISO Press, approved July 2005. http://www.niso.org/kst/reports/standards?step=2&gid=None&project_key%3Austring%3Aiso-8859-1=7cc9b583cb5a62e8c15d3099e0bb46bbae9cf38a

Nunes, José Horta. 2002. "Dicionarização no Brasil: Condições e Processos". In *História do Saber Lexical e Constituição de um Léxico Brasileiro,* edited by Jose Horta Nunes, and Margarida Petter, 99-120. Campinas, S.P: Pontes.

Nunes, José Horta. 2006. *Dicionários no Brasil: Análise e História do Século XVI ao XIX*. Campinas: Pontes.

OCLC Inc. 2023. *Dewey Decimal Classification Glossary.* Dublin, Ohio: OCLC. https://www.oclc.org/support/documentation/glossary/dewey.en.html

San Segundo, Rosa. 1996. *Sistemas de Organización del Conocimiento: la Organización del Conocimiento en las Bibliotecas Españolas.* Madrid: Universidad Carlos III; Boletín Oficial del Estado.

Satija, M. P. 2004. *A Dictionary of Knowledge Organization*. Amritsar: Guru Nanak Dev University.

Soergel, Dagobert. 2009. *Knowledge Organization Systems: Overview,* 1-44. http://www.dsoergel.com/SoergelKOSOverview.pdf

Svenonius, Eliane. 2003. "Design of Controlled Vocabularies." In *Encyclopedia of Library and Information Science*, edited by M.A. Drake. 2nd ed. New York: Marcel Dekker, vol. 2, 822-838.

Wellisch, Hans H. 2000. *Glossary of Terminology in Abstracting, Classification, Indexing and Thesaurus Construction*. 2nd. ed. Medford, N.J.: ASI.

Wersig, Gernot, and Ulrich Neveling. 1976. *Terminology of Documentation: A Selection of 1,200 Basic Terms Published in English, French, German, Russian and Spanish*. Paris: Unesco.

Zang, Marcia Lei, and Philipp Mayr. 2019. "Knowledge Organization Systems (KOS) in the Semantic Web: A Multi-Dimensional Review." *International Journal on Digital Libraries* 20, no.3: 209-30. https://doi.org/10.1007/s00799-018-0241-2