Vol. 51 (2024) No 8

Knowledge Organization

International Journal

Devoted to Concept Theory, Classification, Indexing, and Knowledge Representation

CONTENT

ERGON-VERLAG
ISSN 0943-7444 Knowl.Org.

Knowl. Org. 51(2024)No.8

KO KNOWLEDGE ORGANIZATION

Official Journal of the International Society for Knowledge Organization ISSN 0943 – 7444

International Journal devoted to Concept Theory, Classification, Indexing and Knowledge Representation

Contents

KNOWLEDGE ORGANIZATION KO

Official Journal of the International Society for Knowledge Organization ISSN 0943 – 7444

International Journal devoted to Concept Theory, Classification, Indexing and Knowledge Representation

Ontologies for Semantic Annotation: Proposal for an Ontological Multimedia Reference Model

Daniela Lucas da Silva Lemos* and Renato Rocha Souza**

*Federal University of Espírito Santo (UFES). Av. Fernando Ferrari, 514, Vitória, ES, Brazil

**University of Vienna, Universitätsring 1, 1010 Vienna, Austria

*daniela.l.silva@ufes.br / **renato.rocha.souza@univie.ac.at

Daniela Lucas da Silva Lemos holds an PhD in Information Science from the Federal University of Minas Gerais (UFMG). Professor of the Library Science Department of the Federal University of Espírito Santo (UFES). Professor and Researcher at the Department of Library Science and the Postgraduate Program in Information Science at the Federal University of Espírito Santo (UFES), Vitória, Brazil. She coordinates the Laboratory of Information Science, Data and Technology (Cidat Lab) linked to UFES. Holds a Bachelor in Information Systems Administration, a specialization in Strategic Information Management, a Master Degree in Information Science and a PhD in Information Science. She has experience in the area of Information Science, with an emphasis on Information Organization, Knowledge Representation and Information Retrieval, acting mainly on the following topics: Knowledge Organization, Ontologies, Metadata Standards, Semantic Web, Linked Data and Semantic Annotation of Multimedia Resources on the Web. She has experience with IT, with an emphasis on Databases, Software Engineering and Analysis and Design of Information Systems.

Renato Rocha Souza holds an PhD in Information Science from the Federal University of Minas Gerais (UFMG). Professor of the School of Applied Mathematics of the Getúlio Vargas Foundation (FGV / RJ) and Researcher at the Österreichische Akademie der Wissenschaften. Researcher, Professor and Data Scientist with +20 years of experience in Information Science, Applied Mathematics and Computer Science, working with the topics of Data Science, Machine Learning & Analytics; Scientific Programming, Natural Language Processing; Information Management and Retrieval; Knowledge Organization and Representation; Knowledge Management; Education and Research. Holds a Bachelor in Electrical Engineering, two specializations in Technology and Education, a Master Degree in Production Engineering and a PhD in Information Science. Holds two Post Doctorates in Computer Science, and has international experience as a Visiting Fellow (University of South Wales, UK), Visiting Researcher (Columbia University, NY) and Scientific Researcher (ÖAW, Austria).

Lemos, Daniela Lucas da Silva, and Renato Rocha Souza. 2024. “Ontologies for Semantic Annotation: Proposal for an Ontological Multimedia Reference Model”. *Knowledge Organization* 51, no. 8: 561-581. 45 references. DOI:10.5771/0943-7444-2024-8-561.

Abstract: Recent years have seen considerable growth of online multimedia databases, largely due to digitization processes in different sectors of society. Knowledge organization and representation strategies were used to qualify and enrich data and metadata from different types of documents and ensure persistent and interoperable online information structures. This study aimed to propose an ontological reference model to systematically organize metadata that describes multimedia documents based on different contexts and needs. The proposed model was based on the NeOn methodology and aimed to encompass the functional and nonfunctional requirements for the construction and reuse of ontology classes obtained by merging and aligning previously analyzed multimedia ontologies. This resulted in a comprehensive conceptualization to organize multimedia metadata for application contexts that deal with the semantic annotation of information entities produced and consumed in the web of data (Semantic Web). We concluded that advances in developing conceptual reference models for representing multimedia documents are the result of interdisciplinary efforts that drive progress in the production and use of more consistent and coherent metadata aimed at facilitating the cross-referencing, interconnection and aggregation of online information sources.

Received: 8 September 2024; Revised: 31 October 2024; Accepted 15 November 2024.

Keywords: Knowledge Representation. Conceptual Modeling. Interoperability. Ontologies. Semantic Annotation. Multimedia Metadata Ontology.

1.0 Introduction

In recent years, digitization has played an important role in enhancing the expansion of large online multimedia databases, as is the case in GLAM (Galleries, Libraries, Archives and Museums) institutions (Europeana Tech 2021; Siqueira and Martins 2021; Lemos et al. 2022; Martins et al. 2022). These institutions use modern mechanisms for scanning heritage assets and contemporary information infrastructures, such as digital repositories, in order to democratize scientific and cultural knowledge on the internet.

Collections of multimedia objects, therefore, have grown considerably in their different types, formats, and complexity, including texts, static, and moving images, videos, sounds, three-dimensional (3D) models, websites, and other specific media. Such resources require different forms of processing and representation to link multimedia documents and improve search, browsing, and retrieval systems using aggregating semantic approaches to web resources.

As such, the growth of digital multimedia objects online is considered impractical in terms of their preservation, location, accessibility, interoperability, and reuse without the support of strategies for information and knowledge organization and representation (Martins et al. 2022, 5) that incorporate good digital curatorship practices, such as maintaining, preserving and adding value to data (Higgins 2011).

Svenonius (2000) reported that information must be described in order to be organized, and information representation is the product of this descriptive process. The author highlights that some types of information representation are constructed through languages, subdivided into languages that describe information (content) and those that describe the document (specific media), either as a whole or in parts.

Language used to describe the document is related to descriptive representation (Gilliland 2016; IFLA 2009; Galeffi et al. 2016; Zeng and Qin 2016), also considered a cataloging process, which involves the creation and use of metadata, making it essential in standardizing and describing information resources that give users the ability to find, identify, select, obtain, navigate, and explore the item inside an online catalog (Galeffi et al. 2016). On the other hand, content description language is associated with thematic representation (Lancaster 1986; NISO 2005), which focuses on intellectual and semantic (subjective) aspects such as understanding the subject of the document for the purpose of translation into a documentary language that helps users select search filters and browse online information systems.

Beyond cataloging principles (Galeffi et al. 2016), of which are geared more towards human users, the FAIR data principles (Findable, Accessible, Interoperable, and Reusable) (Wilkinson et al. 2016; Guizzardi 2020) aim to improve the ability of machines to find digital objects and their metadata on the web via a persistent and unique identifier. They also enable access authentication and authorization, harmonic and effective communication with other applications for different purposes, and comprehensive descriptive information for consumption by both humans and computational agents, with a focus on reuse.

In this context, linked open data (LOD) (Bizer et al. 2009; Machado et al. 2019) stands out as a contemporary technique for organizing and processing documents online and involves using W3C open standards to interlink and annotate data. This allows content providers to enrich their metadata schemas with structured and well-defined knowledge specifications based on standards, vocabularies and ontologies, enabling quality information consumption and reuse.

From this perspective, the W3C recommended metadata standards combined with open data principles (Machado et al. 2019) and quality interoperable data (Guizzardi 2020) have been used to organize and represent multimedia information resources, enabling the expansion of access points and improving the management, organization and recovery of online digital objects. However, there are still few advanced studies on the relationship between multimedia information resources and the Semantic Web on this topic (Ferrada et al. 2018; Lemos and Souza 2020).

Research in the fields of Information Science (IS) and Computer Science (CS) have proposed conceptual models based on semantic technologies for reality-based modeling and the search for and retrieval of information in digital environments (IFLA 2009; Galeffi et al. 2016; Charles et al. 2017; Riva et al. 2017; Fink 2018; Lemos and Souza 2020; Guizzardi 2020; Lemos et al. 2022; Bekiari et al. 2024a; Bekiari et al. 2024b) in order to improve the scope of interoperability between different metadata schemas and applications.

Traditionally, the use of metadata is the most common way of adding semantics to documents (Zeng and Qin 2016); however, the Semantic Web proposes annotating document content using domain ontologies (Shadbolt et al. 2006). In the present study, ontologies are viewed as more sophisticated annotation models (Andrews et al. 2012; Lemos and Souza 2020) in terms of semantic data treatment, allowing users to describe and link existing resources through qualifiers such as the concepts, instances, properties, relationships and constraints between these resources.

Ontologies have been used for the semantic annotation of documents in a variety of applications. For example, in archeology, semantic data models are used to document the geometric aspects of fragments of 3D objects that could be reassembled and reconstructed in specific archaeological research (Catalano et al. 2020); in history, ontological conceptual models are applied to code and disseminate data associated with historic photographic archives (Robledano-Arillo et al. 2020); in the field of digital culture heritage, semantic data models are used for the online publication of cultural collec-

tions (Dijkshoorn et al. 2018; Lemos et al. 2022); and in architecture, domain ontologies aim at the reality-based 3D annotation of building conservation state (Messaoudi et al. 2018).

Nevertheless, there are gaps in proposed metadata models for multimedia documents and ontologies for semantic annotation (Van Ossenbruggen et al. 2004; Nack et al. 2005; Lemos and Souza 2020). For example, ISO/IEC standard MPEG-71[1] (Martínez et al. 2002) aims to provide possible solutions for problems associated with producing quality multimedia metadata, but has semantic limitations and the schemas of the different parts that enable descriptions are complex. On the other hand, in terms of proposed ontologies, these generally focus on more generic aspects involving standards, but without considering specific types of metadata for describing multimedia or providing specific descriptors for some types of metadata without relevant modeling reasoning that can ensure, for example, the scalability of their conceptualization. For example, in several of the revised proposals for multimedia ontologies (Lemos and Souza 2020) we expected to find models based on the MPEG-21 Multimedia Framework standard (Kudumakis et al. 2019) for organizing metadata associated with the management of intellectual property rights. This did not occur.

Thus, based on the gaps identified in the above scenario, the present study aims to answer the following opportune and challenging questions: i) *how can a comprehensive conceptual framework that underlies the annotation of multimedia documents be formally expressed?* ii) *what methods and techniques would be suitable for selecting and aligning vocabularies and multimedia ontologies for the annotation of multimedia documents developed by different communities?* and iii) *how can existing types of metadata be systematically organized to annotate multimedia documents for different contexts and needs?*

The present study aimed to advance research on models and modeling for the semantic representation of multimedia documents by proposing an ontological multimedia reference model (OMRM) based on best linked open data practices and on the reuse of existing models in order to cover gaps examined in these, and also expand the coverage of certain important aspects analyzed to systematically organize existing metadata types and describe multimedia documents, according to different contexts and needs. We understand there are advances resulting from the present research for the field of Information Science, especially for the area of Descriptive Cataloging of networked multimedia digital objects, as well as for the area of Ontology Engineering, considering that there are no ontological models for the domain of annotation of multimedia documents that encompass high-level central taxonomic structures, independent of the foundation ontology, and also consider ontological structures based on consolidated ISO metadata standards in library communities digital and multimedia. Additionally, no specific studies were found in the literature that considered categories of descriptive, independent and content-dependent types of metadata represented in reference ontologies of the domain annotation of multimedia documents. Such categories of metadata types are reflected in the way the ontological classes were organized and represented in the OMRM of the present research.

We acknowledge existing ontological models that consider media types and central metadata. However, only for specific domains, such as Europeana's EDM (Charles et al. 2017), IFLA's LRM (Riva et al. 2017; Bekiari et al. 2024b) and ICOM's CIDOC-CRM (Bekiari et al. 2024a), all in the field of Cultural Heritage. The model proposed in this research is intended to be used for multimedia annotations in any domain of knowledge, considering highly relevant multimedia standards such as annotation, decomposition and collection for the realization of the media and the content involved in it.

It is also worth noting that this article is a continuation of research previously published in Lemos and Souza (2020) in which we described in detail the entire methodological path based on the recognized NeOn Methodology, strongly based on reuse on existing models, which led to the careful selection of ontological and non-ontological resources for the OMRM proposed in this article. In this sense, we consider that OMRM is a reference model for the multimedia annotation domain, although it is not yet actually implemented and coded in a computational language, such as Web Ontology Language (OWL) or Resource Description Framework (RDF). However, we consider that the artifact specification delivered in this research already helps in the implementation of OMRM in future research. The OMRM proposed here aims to harmonize (connect) with other vocabularies from any domain available on the network and structured on linked open data principles, including Simple Knowledge Organization System (SKOS), metadata standards with structures in RDF-based languages (e.g.: VRA Core, Dublin Core) and also formal ontologies (e.g.: CIDOC-CRM, DOLCE).

Finally, OMRM's central taxonomic structure, based on a high-level ontology, offers opportunities for harmonization with axiomatized classes that allow structural and semantic alignment with other vocabularies (with or without formal rigor) available for reuse in an open environment.

The above justifications therefore lead us to provide reasons to recommend using the proposed reference model instead of creating an entirely new model.

2.0 Methodology

The proposed OMRM was supported by adopting a current methodological guide, tested and validated for different domains and areas, that follows the guidelines for construct-

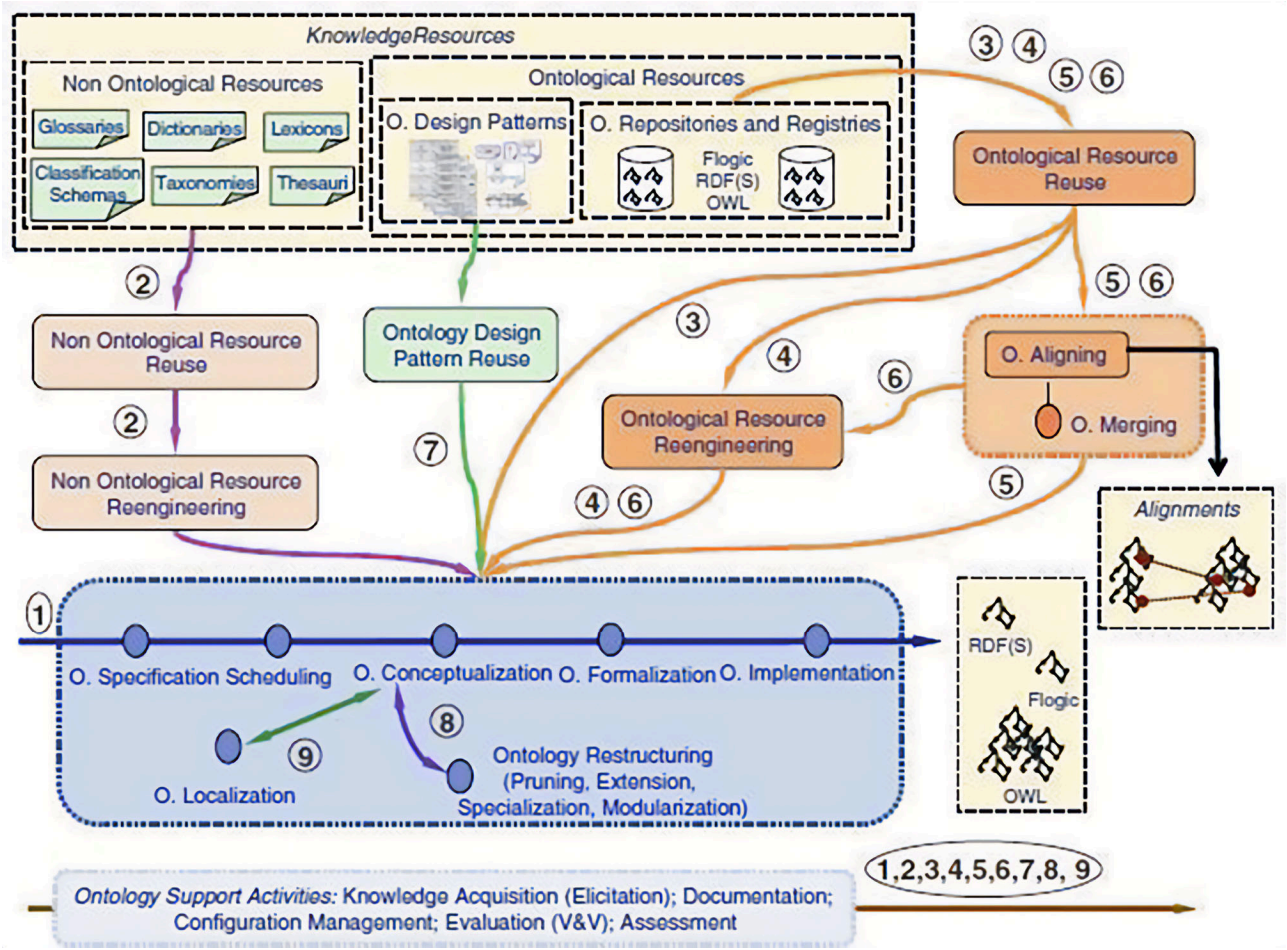


Figure 1. NeOn Methodology Scenario (Suárez-Figueroa et al. 2012, 13).

ing ontology networks based on LOD principles. In this respect, a literature review was carried out in the field of Ontology Engineering and the NeOn Methodology guide was selected from a set of proposals (Silva et al. 2012; Suárez-Figueroa et al. 2012; Falbo 2014; Almeida and Farinelli 2017) because of its LOD practices and the fact that it is the product of methodological frameworks widely accepted in advanced areas such as Software and Knowledge Engineering. Furthermore, some more recent proposals, including the SABiO methodology (Falbo 2014) and the OntoNeo methodology (Almeida and Farinelli 2017), were also derived from the NeOn Methodology.

The NeOn guide covers nine scenarios (Figure 1) that suggest a series of flexible steps for developing ontologies. These scenarios include situations in which ontologies require reengineering, alignment, modularization, localization, support in different languages and cultures, integration with design patterns and non-ontological resources such as metadata standards, dictionaries, thesauri and taxonomies, among others.

Six of the nine scenarios were selected (Scenarios 1, 2, 3, 5, 6 and 8) and are briefly presented in Table 1, which summarizes the results obtained from the methods and techniques applied in the study, including analysis and assessment criteria for ontologies that are candidates for reuse.

The methodological process for these criteria is described in detail in Lemos and Souza (2020, 303-308).

3.0 Results: proposal for an Ontological Multimedia Reference Model

The proposed OMRM was based, a priori, on specifying the requirements for ontologies aimed at multimedia document annotation. The purpose and scope of the model must be defined in order to specify these requirements. The purpose of the model encompasses its intended use, the potential scenarios that require its use and the possible user communities involved in applications that make use of the model, while the scope includes a set of previously determined functional and non-functional requirements.

Stages of the modeling process	Use of NeOn scenarios	Results generated
Identify and select ontological and non-ontological resources in the **multimedia document annotation** domain.	**Scenario 2:** reusing and reengineering non-ontological resources. **Scenario 3:** reusing ontological resources.	**Scenario 2**: **parameter elements** (120 in total), that is, a set of multimedia features based on the MPEG-7 and Dublin Core metadata standards. **Scenario 3**: **ontologies selected** (9 in total) via the literature review and Semantic Web repositories.
Analyze and compare multimedia ontologies according to previously proposed requirements.	**Scenario 3**: reusing ontological resources.	**Weighted ranking** of candidate ontologies for reuse.
Select appropriate multimedia ontologies to reuse knowledge resources for the construction of the proposed model.	**Scenarios 3 and 5**: reusing, aligning and merging ontological resources	**Arrangements** to **organize knowledge resources** according to the types of metadata addressed in the study.
Develop an ontology-based conceptual model for the **multimedia document annotation** domain.	**Scenario 1:** specifying functional and non-functional requirements; and proposing the conceptualization of the ontology. **Scenario 6**: reengineering aligned and merged ontological resources. **Scenario 8:** restructuring the conceptual model to meet the established requirements.	**Class diagrams of the Ontological Multimedia Reference Model**

Table 1. Summary of the results generated by the methodology applied in the study.

The purpose of the OMRM is to represent a consensual conceptualization shared by a given community for the semantic organization of annotations aimed at multimedia documents that are produced, described, published and consumed online, along with their annotations or metadata. Thus, the conceptual model seeks to enrich different types of multimedia metadata via an information framework suited to scenarios that involve, for example, data aggregation in LOD environments, enabling syntactic and semantic interoperability between different institutions and their information systems.

The OMRM can be used, for example, as a key element in the information systems of cultural heritage institutions, whose users consume, interpret, manipulate and generate multimedia content in their collections that are generally digitized and accessible in online digital repositories. The content of these collections archives can be mapped for the OMRM, whose semantic structure incorporates the notion of an Event (a key entity in the cultural domain), making it possible to aggregate actors, objects (physical and abstract), locations and the duration of time intervals. For example, a historical image could be modeled as a sequence of chronological lines containing persistent items (objects and people) combined in events within a time period. This improves the semantic enrichment of data in terms of information retrieval, since data on related (semantically aggregated) events can be collected to create a powerful semantic network of biographical and contextual data on people, documents, objects and places, which would be useful for educational and scientific research. In short, the OMRM makes it possible to classify elements referenced in cultural heritage documents into formal categories, producing legible descriptions of events and objects that improve search, navigation and retrieval systems by aggregating semantic approaches to data from heritage collections.

The OMRM can also be used to expand possibilities for research and collaboration between users in semantically linked knowledge networks on the Semantic Web, including sources of medical, cultural, multimedia, artistic, historical, tourism, educational and social media-related information such as Wikidata, Wikimedia Commons and Wikipedia (Mora-Cantallops et al. 2019; Navarrete and Villaespesa 2021). This would enhance the exploitation of information in integrated knowledge networks, as well as the circulation and collaborative production of information resources that are useful to society. Catalogers are another cat-

egory of users that play an important role in linking annotations in multimedia documents, particularly in dynamic knowledge spaces such as digital repositories and libraries. Finally, but by no means the last possible use for the model, a variety of news websites require efficient methods for organizing multimedia content and transmitting it intelligently to different types of users.

3.1 OMRM requirements

The scope of the OMRM was determined based on functional (FR) and non-functional requirements (NFR) established according to the results of comparative analyses of candidate ontologies for reuse (Lemos and Souza 2020). It is therefore relevant to elucidate FR and NFR within the scope of this study. The former encompasses Software Engineering practices that have been adapted to the field of Ontology Engineering (Fernández-López et al. 1997) in order to facilitate tasks involved in specifying the content of a particular knowledge domain, obtaining ontology-related terminologies. In both fields, NFR are use-related requirements that include performance, usability, reliability, security, availability, maintainability and technologies involved.

It is noteworthy that for FR (1 to 4), competency questions were described that reflect important features of the domain investigated here and any application context that deal with multimedia objects. The competency questions method (Grüninger and Fox 1995) involves determining and applying a series of types and examples of questions formulated in natural or formal (first-order logic) language and empirically designed to be efficiently and correctly answered by the ontology. These questions reflect the main knowledge demands of future users of the ontology. This method enables a practical and intuitive description of the ontology's requirements and scope, helping to obtain a more accurate perspective of the classes, properties and relationships that must be included (Robledano-Arillo et al. 2020). These requirements are described below.

- (FR1) Covers content independent metadata: these data are not directly related to media content, but used to manage and administer information resources, such as creation and production, genre, language, format, usage rights and age-restricted content, among others. Examples of competency questions: i) *where are images created?;* ii) *what age range can access a given program?;* iii) *what are the resolutions of the image files?;* iv) *what are the copyrights of a user related to a certain media (including its content)?*; and v) *which agent is responsible for publishing a copy or part of a work?*
- (FR2) Covers content dependent metadata: the features of visual and audio data are considered primitive or low level in that their content, such as color, texture, shape, spatial relations, movement, location, spectral and temporal timbre and signal parameters, are generally extracted automatically by computer algorithms. Examples of competency questions: i) *what is the predominant color of an image?;* ii) *what part of an audio stream is predominated by the timbre of a musical instrument?; and* iii) *what are the geographic coordinates of an object located in a city shown in an image?*
- (FR3) Covers descriptive metadata: these refer to semantic content that links media entities with their real-world counterparts, such as the face of a person portrayed in an image, as well as aspects involved in personalizing content to facilitate navigation, access and user interaction in relation to content consumption. Examples of competency questions: i) *which frames of a video depict a certain scene?;* ii) *which chapters of a book address a specific subject?;* iii) *which spoken documents portray a particular statement?*
- (FR4) Considers media content and realization in different formats, such as audio, image, text, 3D models and video: separating information objects from their realizations is important in terms of easily visualizing different manifestations (a book in PDF format, a 3D digital replica, songs recorded in an MPEG file) of the objects (a story, a sculpture, a song) and clearly understanding the relationships between them and their realizations. Content-independent metadata such as file size or media location on the web are typically applied to information realization, whereas descriptive metadata for multimedia content aims at describing the message to be conveyed to the consumer of the content. As such, this separation is relevant in that it provides a clear distinction between content semantics and the data itself (e.g. media file). Examples of competency questions: i) *what versions are there for a specific multimedia presentation available online?;* ii) *what are the examples of a literary work in its multiple expressions and manifestations in a bibliographic collection?; and* iii) *what photographs and manuscripts are available in the archival collection for a specific theme, agent or object?*
- **(NFR1) Has an upper ontology as reference:** upper ontologies are referred to as *foundational ontologies* (Guizzardi 2020) and describe very broad concepts such as space, time, matter, objects, events, agents etc. They are considered philosophically well-founded systems of domain-independent categories. Their use semantically benefits the core taxonomy of the domain ontology by clarifying the intended meaning of the terms, supporting, for example, the integration of instances of media content with domain ontologies.
- **(NFR2) Is based on extended multimedia patterns with an LOD approach:** mitigates the challenges of reuse with acceptable and memorable diagrammatic visu-

alizations for a specific set of competency questions (problem and its solution). Some ontologies for multimedia annotation use design patterns to generically organize entities and relationships underlying the multimedia domain, such as annotation and decomposition. These patterns make it possible to link different media resources and coherently integrate metadata (annotations) involved in these resources via a URI, in line with LOD principles. Thus, NFR2 works in conjunction with NFR1 in that the patterns inherit associated axioms and inference services from their upper ontology.

- (**NFR3**) **Considers well-placed ontologies in a ranking** produced from well-founded criteria for reuse: use of a mature, robust and efficient methodology for careful analysis and evaluation of ontologies for multimedia annotation (Lemos and Souza 2020). Reusing available knowledge resources to model knowledge of a domain is recommended in the field of Ontology Engineering (Fernández-López et al. 1997).
- (**NFR4**) **Considers different levels of granularity**: provides a conceptual model that represents a comprehensive taxonomic structure capable of supporting generic (e.g. annotation) and specific multimedia entities (e.g. primitive and specific audio descriptors) according to a particular context.
- **(NFR5) Ensures interoperability** in relation to multimedia content on the web: ensures that the intended meaning of the captured semantics can be shared between different applications within the scope of the Semantic Web. In addition to exchanging multimedia content, the model should also provide the means of transmission in a syntax agreed upon by a community which, in this case, would be via Semantic Web languages such as RDF/OWL.
- **(NFR6) Has an architecture that allows separation of concerns:** provides clear separation of concerns involving the subject of the media (content semantics), knowledge related to managing media information resources (content-independent metadata), structure (media segments) and features of multimedia documents (content-dependent metadata).
- (**NFR7**) **Has an extensible architecture** in terms of building a comprehensive multimedia ontology: since an ontology is always evolving, the inclusion of new concepts must be stipulated in the extensible conceptualization. Extensibility is ensured to the extent that design patterns and upper ontologies can, through meta-categories, expand the possible insertion of new concepts without needing to change the core underlying model.

3.2 Selection, features and alignments of multimedia ontologies suitable for reuse

The comparative analysis of the ontologies for multimedia annotation suitable for reuse was carried out using criteria organized into 4 dimensions (Lemos and Souza 2020, 305-312), as follows:

i) Resource Reuse Effort: estimation of costs related to time and economy required to reuse the evaluated ontology;
ii) Resource Understandability Effort: estimation of effort required to understand the content of the evaluated ontology;
iii) Resource Integration Effort: estimation of efforts undertaken to integrate the evaluated ontology to the new ontology that is being built; and,
iv) Resource Reliability: analysis of the performance of the ontology evaluated against aspects of semantic treatment in declarations (e.g., axioms present, knowledge resources used), evaluation (e.g., available tests) and renowned projects that make use of them.

Weighted ranking (Figure 2) and the findings of comparative analysis of candidate ontologies for reuse enabled the selection and justification of ontological resources according to the previously described requirements. The ontologies best suited to the proposed OMRM were the Media Ontology (1.56), M3O (1.23), COMM (1.19) and M3 Multimedia (0.95). Most of these ontologies use the MPEG-7 metadata standard to support their constituent elements.

The Media Ontology was proposed in 2009 by members of the W3C Media Annotation Working Group, which aims to improve interoperability between metadata schemas for web-based media resources, such as video, audio and images.

The M3O or Multimedia Metadata Ontology (M3O) was created in 2010 as a comprehensive model to represent metadata aimed at multimedia document annotation, including combinations of ontological models commonly used in the Semantic Web.

The COMM or Core Ontology for Multimedia was developed in 2007 by a group of renowned researchers in the fields of multimedia, digital libraries and the Semantic Web. Its main purpose is to provide a sound conceptualization, based on the MPEG-7 metadata standard, that broadly covers a specific domain dealing with multimedia content.

M3 Multimedia was created in 2012 as part of a comprehensive ontology (addressing different domains and languages) denominated the M3 Ontology Network, the product of a Spanish research project involving entities such as the Ontology Engineering Group.

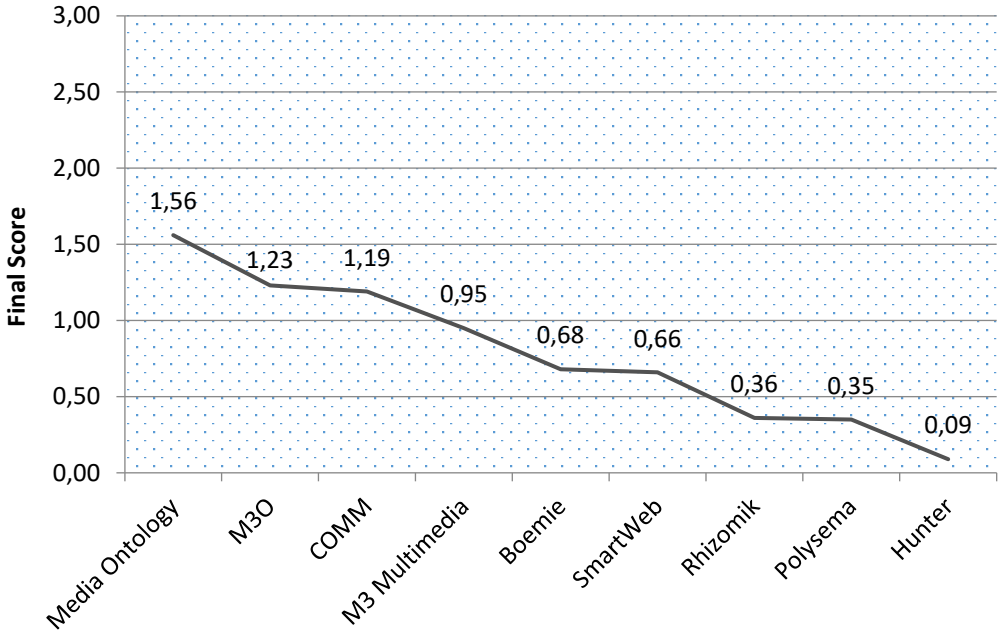


Figure 2. Weighted ranking of candidate ontologies for reuse (Lemos and Souza 2020, 312).

Given that the OMRM must be based on an upper ontology (NFR1) and multimedia design patterns (NFR2), in addition to addressing semantic differences between media content and realization (FR4), M3O was selected as the ontology that best met these requirements and was therefore considered central in the model. This choice is justified by the fact that M3O's conceptualization architecture is based on the upper DOLCE+DnS Ultralight (DUL) ontology (Masolo et al. 2003; Borgo and Masolo 2009) and three design patterns that it references, as described below: Descriptions and Situation (DnS), Information and Realization Pattern and Data Value Pattern.

The *Descriptions and Situation* pattern provides an ontological formalization of context based on role assignment. The semantics embedded in this design pattern (see the application example in the class diagram in Figure 6) state that a situation satisfies (the satisfies relation) a description in which a situation is a specific context with concrete entities that express a certain role; and the description, in turn, is a conceptualization that defines certain concepts which determine (the classifies relation) the roles of entities in a specific context. Entities, on the other hand, are considered relevant or true only in a given context.

Each entity is connected to a given situation via the hasSetting relation, such as quantitative metadata (represented by an entity) related to a color histogram participating (hasSetting relation) in image annotation.

The Information and Realization pattern (Gangemi and Presutti 2009) models the distinction between information objects and information realizations, underpinning FR4. In the example in Figure 3, the class related to the information realization formally realizes some information object (with its inverse is realized by relation). Both concepts are InformationEntity subclasses that allow information to be treated in a general sense (see the application example in the class diagram in Figure 4).

The *Data Value* models the concrete values of an entity in order to reduce the risk of ambiguities (Saathoff and Scherp 2010). DUL (see application example in the class diagram in Figure 4) contains the concepts of *Quality and Region*, which represent, respectively, the intrinsic attributes of an entity linked to its values with corresponding data spaces. In pattern description, the attribute is represented by the concept *Quality*, which is connected to the Entity by the *hasQuality* relation. *The Quality* is connected to a Region via the *hasRegion* relation, and the concrete value is attached to the *Region* by the *hasRegionDataValue* relation. The primitives used in the *Data Value* pattern are useful in expressing structured data values supported by MPEG-7, especially for data automatically extracted from media, such as color, texture and shape, among others.

M3O multimedia patterns are therefore extended from the *Descriptions and Situation* Pattern, including *Annotation Pattern*, *Decomposition Pattern* and *Collection Pattern*, as presented in the next section. Their design diagrams are easily recognizable by the simplicity of their few class and relationship schemes, making it possible to understand the modeling reasoning used in conceptualizations. Addition-

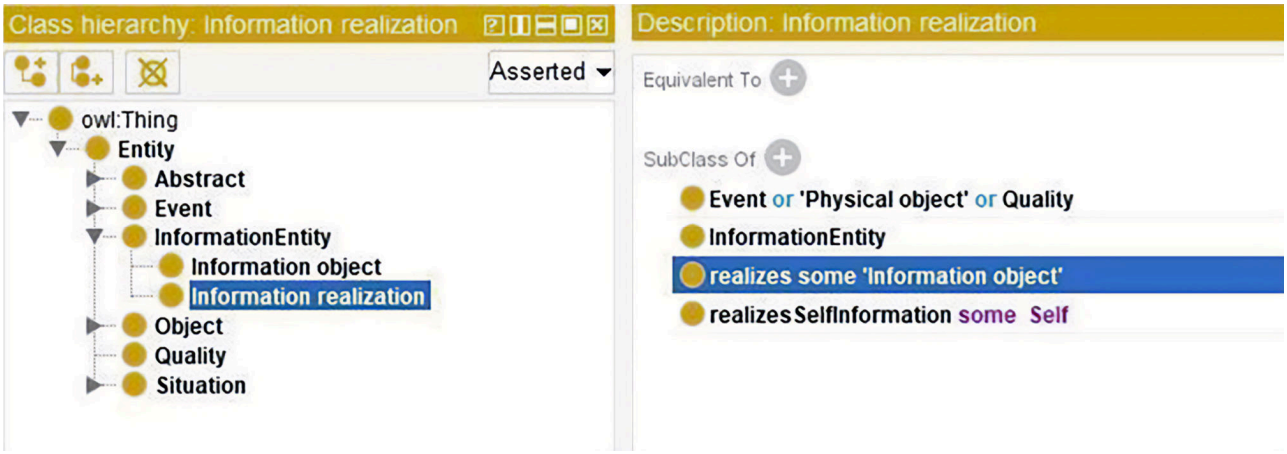


Figure 3. Description of DUL information objects and information realization
Source: screenshot of Protégé software used in the content analysis of the ontology DUL.

ally, all three multimedia patterns (annotation, decomposition and collection) act on the semantics specified in the *Information and Realization* pattern, enabling objects of information and their realizations to be annotated, decomposed and organized into collections.

Media Ontology is recommended for content-independent metadata (FR1) because of its satisfactory coverage index (Lemos and Souza 2020, 313) for this type of metadata (in relation to the other ontologies analyzed), particularly for descriptors aligned with the Dublin Core metadata standard.

The COMM and M3 Multimedia ontologies provide descriptors that align well with content-dependent metadata (FR2), useful in computer processing of digital data for the automatic generation of quantitative metadata. Both have very similar visual coverage indices (Lemos and Souza 2020, 310-311), particularly for descriptors involving color, texture, shape and location of regions of interest. Metadata to describe 3D characteristics, for example, are present in both ontologies, primarily for shape-related visual aspects, since both are based on MPEG-7 for multimedia content description. MPEG-7 includes descriptors for the geometric characteristics of 3D objects, such as symmetry, circularity, axis location, size and orientation of consecutive border segments, curvature points and angles of curves. Knowledge resources related to audio metadata can be selected from M3 Multimedia because it reuses both the visual and audio metadata from the VDO Boemie ontology (Lemos and Souza 2020).

Descriptive metadata (FR3) aimed at the semantics of media content are generally linked to instances of domain ontologies or controlled vocabularies with less formal rigor (called Simple Knowledge Organization System - SKOS) whose semantic labeling is organized within the taxonomy of an upper ontology. Since M3O is part of DUL, it plays the role of organizing semantic labels from domain ontologies or SKOS into abstracts entities such as event, object, time, place, etc., in addition to dealing with their relationships. M3 Multimedia covers properties for navigation and access (content customization), audio descriptors with high-level features (spoken content, for example), and common descriptors for segment annotation.

It should also be noted that because M3O uses an upper ontology as reference and is based on multimedia patterns extended from ontology design patterns, all its features meet the previously outlined nonfunctional requirements, such as interlinking open license data (media and its metadata), treating different levels of granularity, interoperability, separation of interests and extensibility. The possibility of linking different media resources and integrating metadata is achieved by a semantic URI that uniquely identifies the resources (entity at its most abstract level) in the network. The data value corresponding to the URI can be modeled through the Entity in DUL (rdfs: domain primitive) by setting any URI (rdfs: range primitive).

3.3 Arrangements and mappings of the OMRM ontology classes

Following the alignment of ontologies with the predetermined requirements, as elucidated in the previous section, knowledge resources were semantically organized into proposed arrangements or groups, including: i) information objects (document content) and their realizations (media) involved in the annotation context; and ii) types of multimedia metadata (particularly from MPEG-7) and their respective ontology classes based on content-dependent, content-independent and descriptive metadata. That said, the

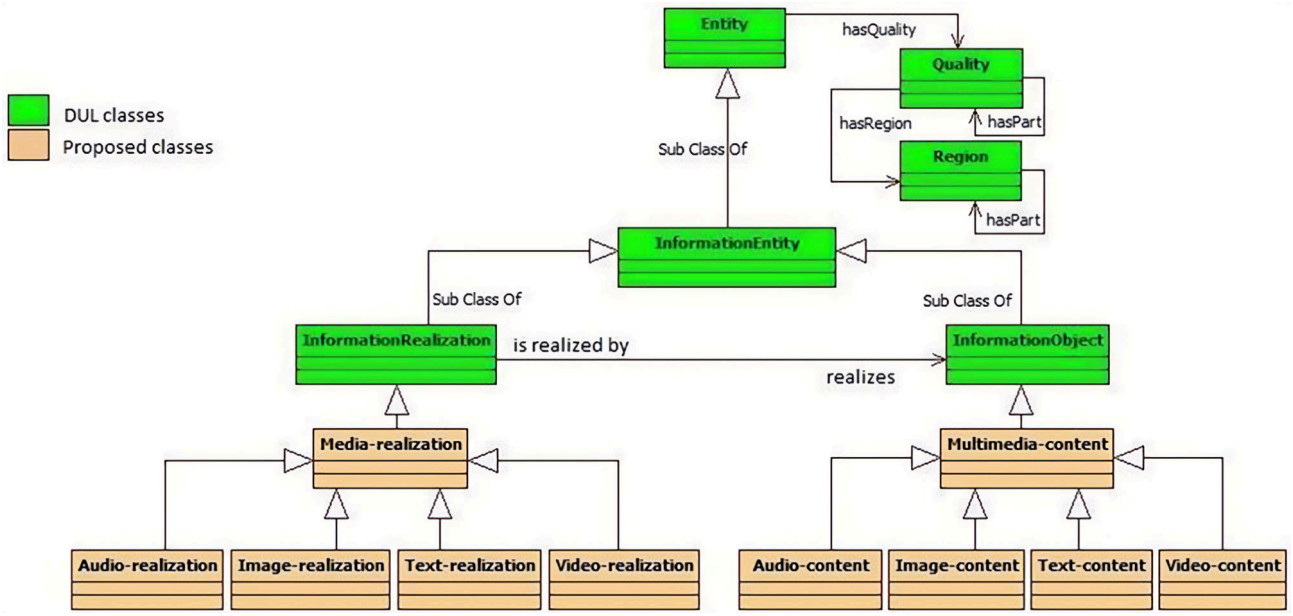


Figure 4. Classes of key entities in the Ontological Multimedia Reference Model.

mappings were determined from the arrangements and multimedia patterns proposed for the model.

In the first arrangement (Figure 4), information objects and their realizations were grouped into the *Multimedia-content* and *Media-realization* ontology classes, respectively. These classes were then generalized to the DUL classes *Information object* and *Information realization*, respectively.

New media specializations can be inserted into the OMRM taxonomy for specific applications, given the expansible nature of the model.

Tables 2 and 3 present the remaining arrangements organized into hierarchical structures for organizing knowledge resources according to the types of multimedia metadata addressed in the study. In spite of these tables presenting informal definitions, they are considered important tools to support communication, in addition to enabling the discussion, negotiation and representation of the consensus reached among domain specialists for future implementations involving specific domains that deal with multimedia metadata.

We opted to map the taxonomic structure of the ontology by identifying its subclasses and superclasses. The "." symbol (point) indicates the subclass relationship between the classes involved; for example, color as a subclass of visual metadata. Corresponding ontology classes were mapped (or proposed) for each ontology involved in reuse and for some concepts from MPEG-7.

It is important to underscore that the nomenclature of the classes was maintained according to their ontological origin, whereas the proposed classes were named in line with naming conventions underlying the use cases studied. The (*) symbol signals a new class for the model.

Ontology models based on design patterns (such as COMM and M3O) already contain formal coherent groupings (provided by the axioms of their upper ontologies) represented in specific ontology classes. For example, the COMM *localization-descriptor-parameter* class contains the concepts *region-locator-descriptor-parameter, bounding-box* and *region-boundary*, which correspond to metadata for visual localizations.

On the other hand, the properties of ontology models focused on relations and attributes (such as Media Ontology and M3 Multimedia) are grouped into ontology classes. In these cases, axiom modeling is necessary in order to formally establish the metadata elements of the groupings as a specific type. Constraints can be modeled using logical statements including existential quantifiers to indicate relationship to at least one individual, and universal to indicate relationship to all individuals. A constraint could be modeled, for example, for the Media_Creation class (shown in Table 2) declaring that there is a unique identification (class identification) and a location where the resource can be accessed (class locator) associated with an instance of the entity class with media creation and production roles. Figure 5 displays a code example representing such an axiom.

Other types of axioms can be created, for example, to restrict the participation of metadata in the annotation of an information entity, considering the semantic distinction between an information object and an information realization, subjects of the annotation. In the example in Figure 5, the locator metadata from the Media Ontology has the role

Content-Independent Metadata				
	Media Ontology			
		Type of metadata	**Ontology Class**	**Class Description**
		Media creation and production	Media_Creation (*)	Describes the features involved in creating media content and its associated resources.
		Media classification	Media_Classification (*)	Describes features aimed at classifying media, such as genre, subject, purpose and language, among others.
		Media information	Media_Information (*)	Describes types of storage, including content format, compression and coding.
		Media usage	Media_Usage (*)	Describes features that reflect the usage rights, registration and availability of media usage.
Content-Dependent Metadata				
	COMM Ontology			
		Type of metadata	**Ontology Class**	**Class Description**
		Visual	structured-data-parameter.visual-descriptor-parameter	Describes primitive visual features for color, texture, shape and motion.
		Color	.color- descriptor-parameter	Describes various descriptors and supporting parameters in the representation of different aspects of color-related features.
		Texture	.texture-descriptor-parameter	Describes important aspects in revealing tactile, depth and surface orientation features for an image.
		Shape	.shape-descriptor-parameter	Describes features related to the spatial arrangement of pixels that belong to an object or region. The descriptors can be grouped into 2D or 3D classes.
		Motion	.motion-descriptor-parameter	Describes spatial and temporal features captured by camera movement, a moving object, or both.
		Localization	localization-descriptor-parameter	Describes localization for regions of interest in spatial and spatiotemporal domains.
	M3 Multimedia Ontology			
		Type of metadata	**Ontology Class**	**Class Description**
		Audio	LL_Audio_Descriptor	Describes primitive descriptors involving spectral, parametric and temporal features to describe audio signals and files.
		Spectral basis	.Spectral_Basis_Descriptor	Describes low-dimensional projections of a high-dimension spectral space to aid in compactness and identification.
		Spectral timbre	.Spectral_Timbral_Descriptor	Describes timbre features related to the signal spectrum.
		Temporal timbre	.Temporal_Timbral_ Descriptor	Describes temporal features of audio segments; particularly useful in describing the timbre features of musical instruments.
		Signal parameters	.Signal_Parameter_ Descriptor	Describes periodic or quasi-periodic signals.
		Basic spectral	.Basic_Spectral_Descriptor	Describes descriptors derived from signal frequency analysis.
		Basic	.Basic_Descriptor	Describes basic descriptors for general use and applicable to all types of signals.

Table 2. Arrangements for types of content-dependent and content-independent metadata.

of annotating media files that are located on the Web. Such an assignment should only apply to one realization of the information and, therefore, should be formally declared in the locator annotation class. On the other hand, the description metadata (also from the Media Ontology) has the role of annotating media content information. Thus, this metadata should only apply to information objects, and therefore an axiom should be created for the description annotation class to enforce such a restriction.

Following the proposed arrangements involving the knowledge resources of the study (indicated in Tables 2 and 3), the ontology classes of the M3O multimedia design pat-

Descriptive Metadata		
M3 Multimedia Ontology		
Type of metadata	**Ontology Class**	**Class Description**
Navigation and access	Navigation_Access (*)	Describes aspects of features that facilitate navigation and access to multimedia content, such as summaries.
High-level audio	HL_Audio_Descriptor	Canonically describes a sound with a certain degree of generality, including descriptors aimed at covering specific domains.
Spoken content	.Spoken_Content_ Descriptor	Describes details of spoken words in an audio stream.
M3O Ontology		
Type of metadata	**Ontology Class**	**Class Description**
Organization of digital objects into collections	CollectionPattern	Describes features of collections of information entities with common properties.
Media segments	DecompositionPattern	Describes the structure of multimedia content in terms of segments, such as frames, moving and static regions and audio tracks.
Content semantics	DUL:Entity	Describes real-world objects, events and notions that can be abstracted from multimedia content.
MPEG-7 metadata standard		
Type of metadata	**Ontology Class**	**Class Description**
Temporal segment	Temporal_Segment (*)	Describes a set of temporal features related to segment decomposition for specific media content, such as video, audio, scenes and moving regions.
Spatial segment	Spatial_Segment (*)	Describes a set of spatial features related to segment decomposition for specific media content, such as 2D and 3D images and moving regions.
Spatiotemporal segment	Spatio_Temporal_ Segment (*)	Describes a set of spatiotemporal features related to segment decomposition for specific media content, such as moving and audiovisual regions.

Table 3. Arrangements for types of descriptive metadata.

terns (*Annotation Pattern*, *Decomposition Pattern* and *Collection Pattern*) were semantically mapped, as shown in the class diagrams described below.

In order to ensure better visualization and understanding, the OMRM was segmented into three parts associated with the multimedia design patterns underlying the proposed conceptualization. For the purpose of easy visualization, the diagrams depict more general as opposed to specific classes.

In M3O, an *AnnotatedConcept* classifies an *InformationEntity* that is the information resource (physical or digital object) to be annotated (as a whole or in parts). Each metadata element is represented by an *Entity* (with a semantic URI) classified by an *AnnotationConcept*. The mappings resulting from this conceptual framework model an information entity (which may be an information object or information realization) and the metadata that participate in the annotation process.

The classes (derived from groupings or mapped from the corresponding ontologies) referring to multimedia metadata were specialized (*Subclass Of* relation) in the *AnnotationConcept* class (Figure 6), which assigns the data entities the role of annotation and formally describes their nature as metadata. For example, an image object (information object) that requires annotation by a semantic concept from a domain ontology (eg.Wikidata) would be classified by an *AnnotatedConcept*. The domain ontology instance that plays the role of semantic metadata would be classified as an *AnnotationConcept*. The link between the image object and the instance of a Wikidata semantic structure is established by the *hasSetting* relation, whereby, as a rule, all the DUL *Entities* (event, object, agent, place and time) has a '*hasSetting'* with the annotation situation that satisfies the *AnnotationPattern*.

The class diagram in Figure 6 shows semantic mapping for the previously described annotation pattern.

Figure 7 depicts the taxonomic structure of the annotation pattern, indicating the axiomatization of the AnnotationSituation class.

Mappings involving the Decomposition Pattern are ensured by the media type classes aligned with the DUL information entities, as shown in Figure 4. As such, a *Composite-*

```
<owl:Class  rdf:ID="Media_Creation">
    <rdfs:subClassOf>
      <owl:Restriction>
        <owl:someValuesFrom>
            <owl:Class  rdf:ID="identifier"/>
               </owl:someValuesFrom>
                 <owl:onProperty>
              <owl:ObjectProperty  rdf:resource="&DUL;'has  part'"/>
      </owl:Restriction>
    </rdfs:subClassOf>
    <rdfs:subClassOf>
        <owl:Restriction>
          <owl:onProperty>
            <owl:ObjectProperty  rdf:resource="&DUL;'has  part'"/>
               <owl:someValuesFrom>
                 <owl:Class  rdf:ID="locator"/>
               </owl:someValuesFrom>
        </owl:Restriction>
    </rdfs:subClassOf>
</owl:Class>
```

Figure 5. Example of restriction for the Media_Creation class of the proposed model.

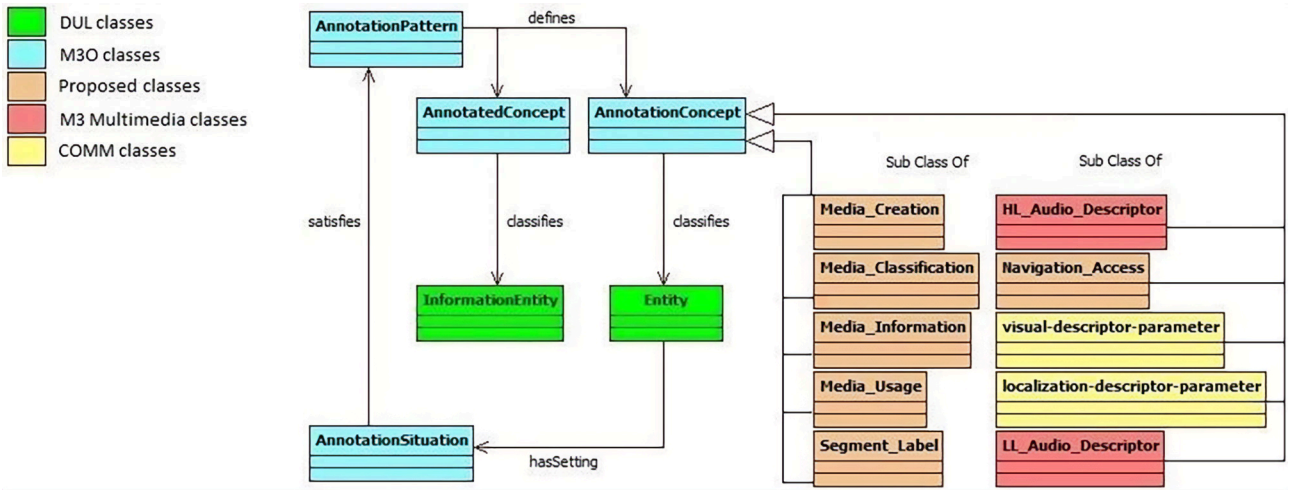


Figure 6. Classes of annotation entities in the Ontological Multimedia Reference Model.

Concept plays the role of an object or document involved in decomposition and the *ComponentConcept* that of the segments resulting from its decomposition. Both classes are configured with a semantic URI inherited from the DUL *Entity class*.

In the case of *CompositeConcept*, the use of the *owl:disjointWith* constructor to limit the participation of instances of media types in inappropriate classes is crucial. For example, an *Audio-content* is disjoint from *Video-content* and *Image-content*. For *ComponentConcept*, the classes proposed for segment types resulting from decomposition were specialized (*Sub Class Of* relation) as *ComponentConcept* subclasses named *TemporalSegment*, *SpatialSegment* and *SpatioTemporalSegment*, representing temporal, spatial and spatiotemporal features of dimensions, respectively.

For these classes, axiom modeling is important in order to impose restrictions on the segment types that form valid decompositions for content involving specific media. For example, a segment corresponding to a moving region (specialized from the *SpatioTemporalSegment* class) would be classified only as video media, and its localizations aimed only at annotation classes involving metadata on visual and time localizations.

This highlights the participation of the annotation pattern in describing the resulting segments for the type of metadata involved, including information on the access location, creator, media, and usage license of the segment.

The class diagram in Figure 8 shows semantic mapping for the previously described decomposition pattern.

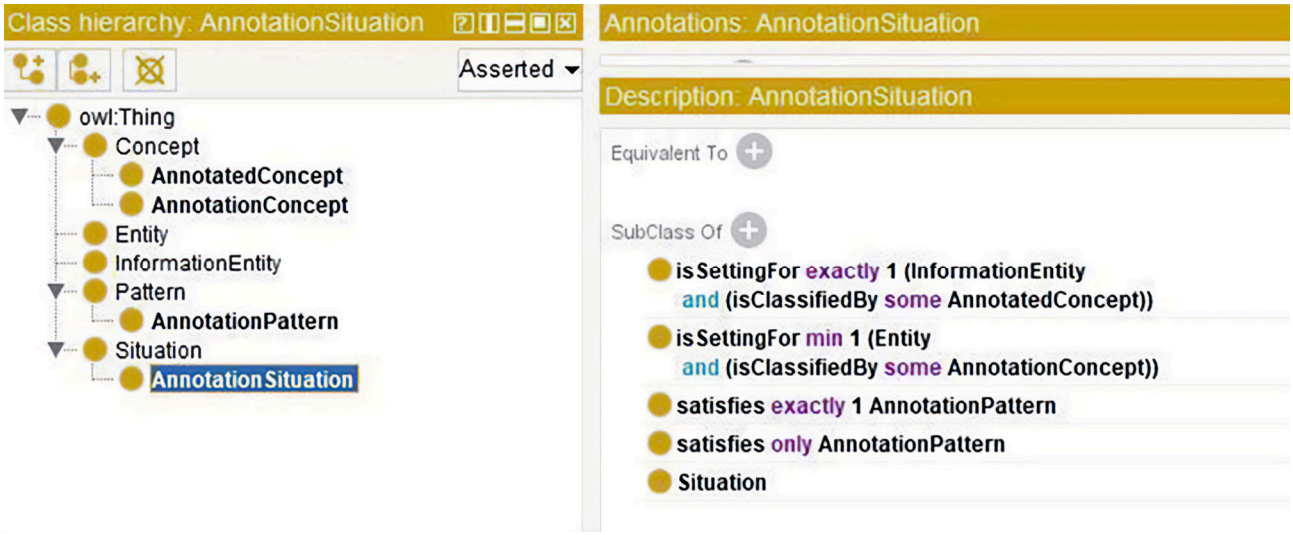


Figure 7. Annotation of information entities.
Source: screenshot of Protégé software used in the content analysis of the ontology M3O.

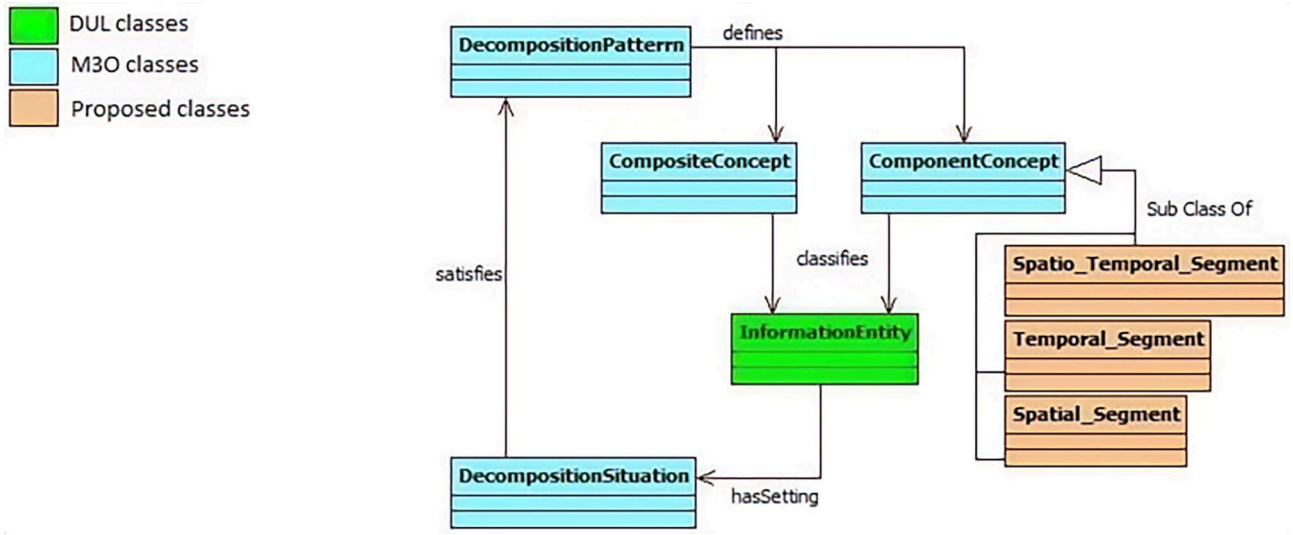


Figure 8. Classes of decomposition entities in the Ontological Multimedia Reference Model.

Figure 9 shows the taxonomic structure of the decomposition pattern, indicating the axiomatization of the *DecompositionSituation* class.

The multimedia pattern for the M3O collection makes it possible to represent collections of information entities with common properties via *CollectionPattern*, which supports the collaborative creation of collections by taking the source or origin of the information entities involved into account. A set of images collected by different people about a common subject is an example of a collection, which can be shared on an online community such as *Flickr*.

The core concepts of the collection pattern establish specializations with DUL design patterns. The *CollectionPattern* class stipulates the existence of exactly one *CollectionConcept* that classifies an *InformationEntityCollection*, which, in turn, is a collection of information entities. Finally, *AnnotationPattern* is integrated and interacts with collection pattern classes in that it provides classes of multimedia metadata to annotate entities in *InformationEntityCollection*.

The class diagram in Figure 10 shows semantic mapping for the previously described collection pattern.

Figure 11 depicts the taxonomic structure of the collection pattern, indicating the axiomatization of the *CollectionSituation* class.

After semantic mapping, the following tasks are recommended, based on the ontology engineering methodologies

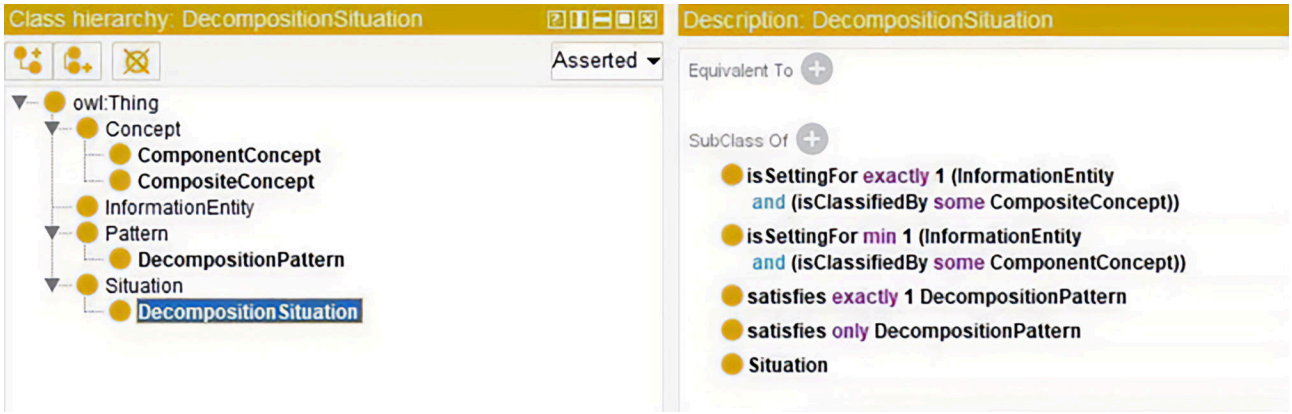


Figure 9. Decomposition of information entities.
Source: screenshot of Protégé software used in the content analysis of the ontology M3O.

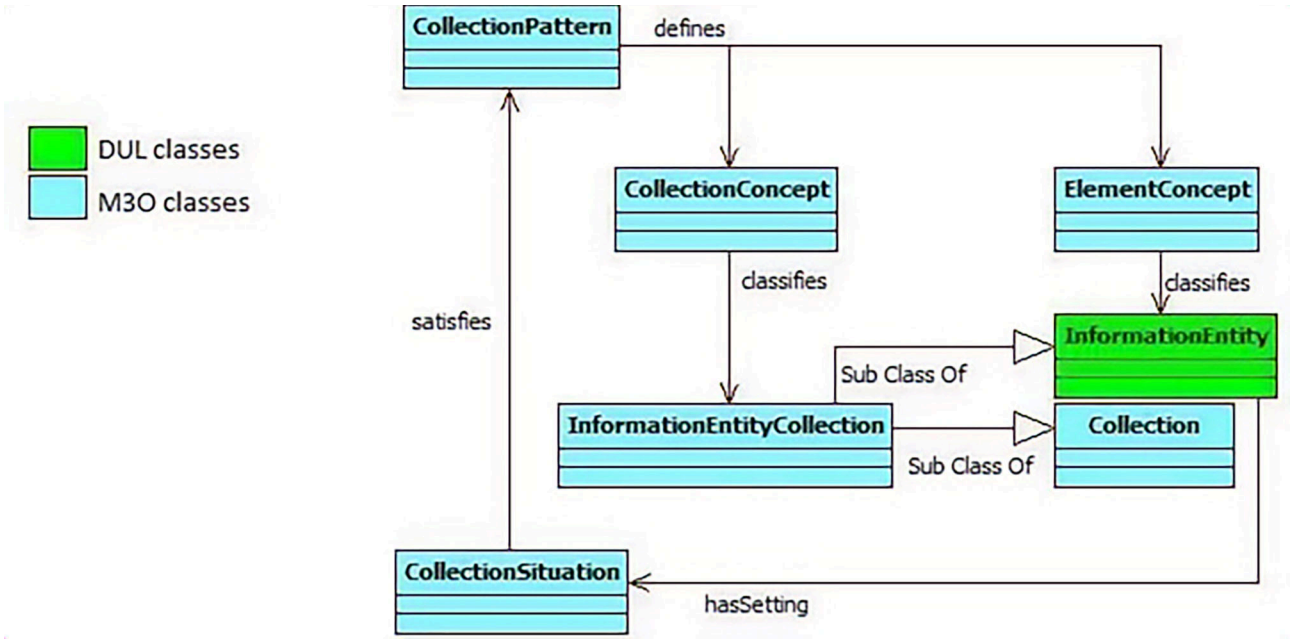


Figure 10. Classes of collection entities in the Ontological Multimedia Reference Model.

operationalized during the study: i) remove unnecessary concepts from the conceptualization of the resulting ontology to prevent an extensive taxonomy with ambiguous concepts; ii) clearly and accurately document all the ontological elements in the conceptual reference model; and iii) validate the taxonomy to assess the consistency of the resulting ontology.

Finally, experts who deal with multimedia files should be consulted to ensure that the domain can be modeled satisfactorily. For example, Library and Information Science professionals are experts in descriptive and subject cataloging of information resources and can contribute by establishing agreements on specific metadata for each information entity involved in model specification. Computer vision, image processing and audio signal experts can contribute to modeling decisions regarding content-dependent metadata, which are heavily dependent on technical knowledge in these areas. This allows ontology engineers to focus on tasks related to modeling domain knowledge.

4.0 Discussion

The proposed ontological model can be characterized as a reference in conceptual specification for multimedia documents that specifically target internal curatorship in information systems aimed at designing normalized and enriched open databases, with a view to improving contemporary information retrieval systems (Mora-Cantallops et al. 2019;

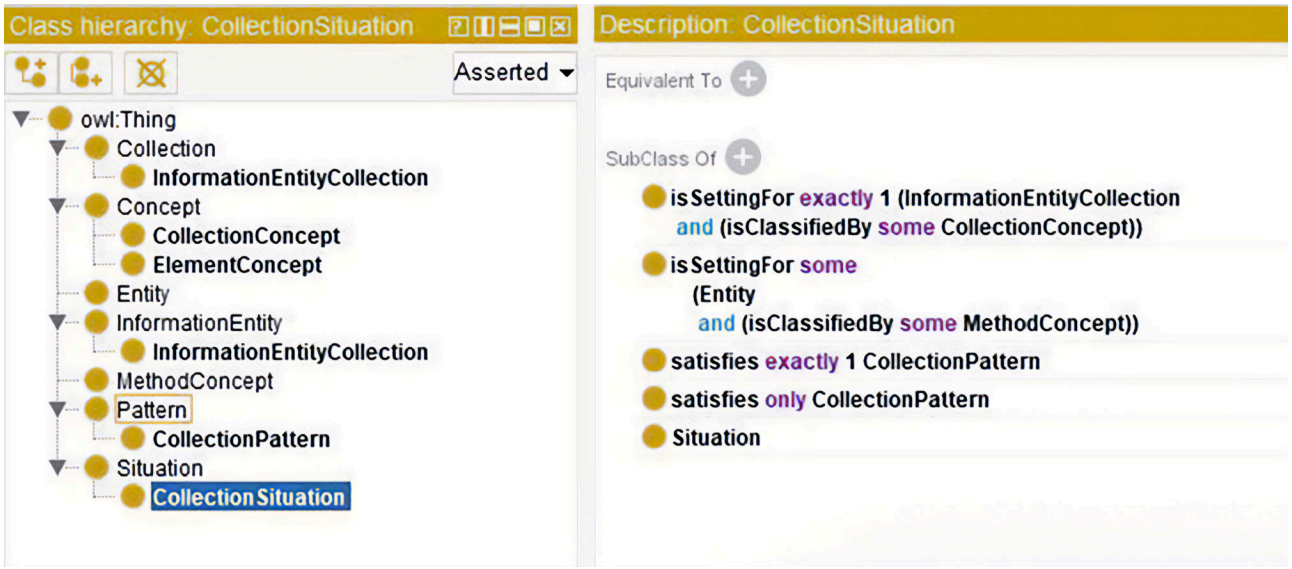


Figure 11. Core concepts of the collection pattern.
Source: screenshot of Protégé software used in the content analysis of the ontology M3O).

Navarrete and Villaespesa 2021; Siqueira and Martins 2021) by providing useful metadata that enable the discovery, reuse, aggregation and integrated search for multimedia objects online.

Metadata standards can be considered the product of the historical development of bibliographic standards and are therefore linked to cataloging codes, conceptual models and new methodological elements for managing information, such as Resource Description and Access (RDA) and FAIR and LOD principles in different online media (Zeng and Qin 2016; Martins et al. 2022). As such, they play a vital role in describing information resources, resulting in qualified access points for discovery, search and retrieval.

However, over the course of the study, significant problems were identified in relation to multimedia metadata (Van Ossenbruggen et al. 2004; Lemos and Souza 2020), mainly due to the convergence of information processes on the web, namely: i) cost: producing quality interoperable and linkable metadata is a costly and time-consuming process; ii) subjectivity: human annotators generally have specific views about content and the context in which it is used; iii) restrictiveness: a schema with few restrictions (such as free text fields) generally provides subjective and inconsistent terminology that is not easily machine readable; iv) longevity: constructing an annotation schema for specific purposes that is sufficiently generic to encompass different domains is a difficult task; v) privacy: metadata can contain private or confidential information, which requires special care; and vi) standardization: there is a need for syntactic and semantic-level standardization to achieve interoperability between different metadata schema and applications.

Based on these challenges, metadata as products and processes evidently require specific modeling aspects. Despite the comprehensive and commonly used definition of metadata found in the information science literature, its uses, syntaxes and applications differ in scale, complexity and cost. The results of this study contribute by providing possible solutions on how to efficiently index, catalog and retrieve multimedia content considering the numerous types of existing metadata for different needs and situations, as explained below.

The OMRM aimed to cover the functional characteristics established for multimedia metadata (content-independent, content-dependent and descriptive) aimed at representing these types of documents.

The *Media_Creation*, *Media_Classification*, *Media_Information* and *Media_Usage* classes are intended to ontologically organize metadata related to the management of information resources (content-independent – FR1), which can be applied to the realization of media and its content, thereby promoting the organization of high-level descriptions of different domains that require this semantic distinction to represent and retrieve information resources. Such classes were modeled based on a proposition of reuse of the Media Ontology (Lemos and Souza 2020), an ontology proposed to define a set of central annotation properties to describe multimedia content, along with a set of mappings between the main metadata formats currently in use, such as EXIF, IPTC, DIG35, Dublin Core, MPEG-7, among others. For example, the models used in Cultural Heritage Domains (e.g.: the Functional Requirements for Bibliographic Records - FRBR - Family of Models, or the

Library Reference Model - LRM) (IFLA 2009; Riva et al. 2017; Bekiari et al. 2024b) does not cover administrative metadata important for bibliographic universe (Riva et al. 2017, 15), such as copyright (covered by *Media_Usage* class from OMRM model), preservation metadata or acquisition processes (covered by Media_Creation and Media_Information classes from OMRM model). We can also mention the need of semantically formalized (via axioms) information structures for different types of media (e.g.: a video collection, a music album, works of art, films, books), allowing content providers, such as digital libraries or museums, to identify and interpret the different copyrights associated with their multimedia content resources. In this sense, the aforementioned ontological classes of the OMRM model would allow, due to: i) the levels of extensibility based on the high-level DUL ontology; ii) the integration with complementary MPEG ISO/IEC[2] initiatives - such as the MPEG-21 Multimedia Framework; and iii) its set of standardized ontologies for the codification of media-related intellectual property rights information (Kudumakis et al. 2019); would allow for the transparent use of multimedia services across a wide range of networks and devices for diverse users.

The *visual-descriptor-parameter*, *localization-descriptor-parameter* and *LL_Audio_Descriptor* classes are aimed at the ontological organization of quantitative metadata (content-dependent – FR2) for visual aspects and localization in spatial, temporal and spatiotemporal regions, as well as audio in media content. For these types of metadata, the *Data Value* pattern of the DUL upper ontology, including relations between a *Quality*, *Region* and *Entity*, represent ambiguity-free features and data values for annotations of this nature. This paves the way, for example, for processing documents from a complex domain such as tangible cultural heritage (e.g.: archeological artifacts, sculptures, buildings), which typically involves the segmentation of physical or digitized structures. Thus, semantic annotation for digital measurements of the resulting 3D fragments (Catalano et al. 2020) creates reliable counterparts for the reassembly of the complete artifact. This favors the integrated search for and retrieval of heterogeneous federated data in the field of cultural heritage.

The *Navigation_Access* and *HL_Audio_Descriptor* classes (FR3) were proposed to ontologically and sequentially organize navigation and access to content personalized according to users' media preference (such as a movie synopsis), and high-level audio features, including a descriptor for indexing spoken content in audio streams (Martínez et al. 2002). Examples of applications include: i) a film or video recording in which a character speaks a particular word or sequence of words; the media support would then be recognized and the query would return content at the specific media position; ii) databases of spoken documents that enable the position of discourse in corresponding audio documents to be retrieved; and iii) retrieving a photograph annotated by a statement.

The classes related to the *Collection* and *Decomposition Patterns* (*Spatio_Temporal_Segment*, T*emporal_Segment* and *Spatial_Segment*) promoted the organization of multimedia collections and multimedia segments, respectively. Both are related to the *Annotation Pattern* which, in conjunction with its specialized classes of metadata types (described above), assign semantic-level links to and between media using LOD principles (Bizer et al. 2009) via specific annotation in content or in media realization (FR4).Structural annotations describe the structure of multimedia content in terms of video segments (frames, moving and static regions) and audio segments, while annotations on content describe real-world objects, agents, events and notions (FR3) that can be abstracted from multimedia content and linked to knowledge organization systems (Lemos and Souza 2020; Lemos et al. 2022) such as ontologies and SKOS of specific domains available online.

In the field of descriptive cataloging, the absence of semantic standards to describe multimedia digital objects at the levels of data structure, value, content and communication (Martins et al. 2022, 7) causes serious problems in standardization, normalization, quality and exchange of descriptions in a linked open data environment (Machado et al. 2019) that could be remedied by Semantic Web vocabularies that align with the by cataloging principles (Galeffi et al. 2016), for example the RDA Element Sets[3] and the RDA Value Vocabularies[4] that were created from attributes and relationships defined in Resource Description and Access (RDA[5]) at the RDA registry.

In this context, OMRM can contribute to the field of descriptive cataloging and its high-level principles focused on the user's tasks of finding, identifying, selecting, retrieving, navigating and exploring the item within an online catalog as a search and discovery system on the web. Therefore, the reference model seeks to portray generic and specific aspects of the multimedia document (at various levels of multimedia entities granularity) that make it unique by establishing well-founded semantic (with the use of a upper ontology) and interoperable access points that allow users greater search and retrieval capabilities in the web of data, making relationships explicit and providing contextual information.

Additionally, the OMRM can contribute to the FAIR guiding principles (Wilkinson et al. 2016) by allowing the creation of consistent and persistent identifiers (e.g.: URI and IRI, among others) of multimedia objects (data) and their annotations (metadata), whose formal semantics can be used in different online datasets for navigation, collection, extraction, mapping, enrichment, aggregation and other possible human or corresponding automated actions that favor the localization ('F'), access ('A'), interoperability

('I') and reuse ('R') functions. As reported by Guizzardi (2020), the "I" (interoperability) of FAIR is only possible with the support of information structures (e.g.: metadata standards, controlled vocabularies, cataloging rules and usage licenses) that are ontologically consistent and explain the ontological commitments that they make. According to the author, the description of real-world objects requires more than vocabularies, but the use and reuse of good domain ontologies.

As such, data providers that use a FAIR-based data quality policy, for example, could benefit from an aggregation service for online collections of multimedia objects (as a whole or in parts) to expand the search possibilities of their users in knowledge networks semantically linked to the web of data.

Problems associated with the syntactic and semantic interoperability requirements of online multimedia applications can be mitigated by the formal nature of the DUL upper ontology and its design patterns for ontology content. These structures ensure that the intended meaning of the captured semantics in the reference model can be shared between different applications within the scope of the Semantic Web.

5.0 Conclusion and future work

The results obtained in this study reflect numerous contributions to the fields of information and technology, particularly for knowledge and information organization and representation. Research in this area focuses largely on document processing (cataloging, indexing and classification), especially issues linked to the semantic nature of information. It should be noted that the present study used document processing concepts, theories, principles and methods, including content analysis, cataloging, classification, categorization and conceptual modeling. These practices supported the understanding, interpretation and systematization of knowledge resource content (metadata standards, controlled vocabularies, conceptual models and ontologies) in the OMRM architecture.

The central framework of the OMRM is a conceptual ontology based on cognitive, philosophical and linguistic aspects that provide metacategories to formally describe events, objects, time, space and others in order to semantically organize content from domain-specific ontologies. In this respect, formal semantics from the OWL representation language contribute considerably to the scope of the proposed conceptual structure that aims to describe any aspect related to multimedia data.

Comprehensiveness, in turn, is achieved via the application of Ontology Engineering principles, which suggests the use of upper ontologies and design patterns for ontology content. Thus, the OMRM seeks to ensure its connection with metadata, controlled vocabularies, SKOS and domain-specific ontologies via axiomatized definitions of high-level concepts from the DOLCE+DnS Ultralight ontology and its *Description and Situation*, *Information and Realization* and *Data Value* design patterns, used to generically organize entities associated with multimedia content, such as annotation, decomposition and collection. The central taxonomic structure of the proposed model therefore provides a solution to the first research question regarding *how to formally express a comprehensive conceptual framework that underlies the annotation domain for multimedia documents*.

The second question, *what methods and techniques would be suitable for selecting and aligning vocabularies and multimedia ontologies for the annotation of multimedia documents developed by different communities?*, was solved by using the NeOn Methodology guide, which is based on LOD initiatives for the construction of network ontologies. The methodology proved to be robust and efficient at explaining the different dimensions and variability in the analysis of knowledge resources identified in the literature and Semantic Web repositories, which ensured the reuse of ontologies suited to the conceptualization of the OMRM.

The third question, regarding *how to systematically organize existing types of metadata to annotate multimedia documents for different contexts and needs* was solved by covering functional requirements from metadata types (content-independent, content-dependent and descriptive) modeled here using relevant modeling reasoning that ensures the extensibility of the conceptualization to annotations (in both the realization of media and its content) aimed at different contexts and scenarios.

Thus, the objective of the study was achieved and contributes to the well-grounded proposal of an ontological conceptual model for the semantic organization of multimedia metadata for different application contexts that publish and consume data on the web.

It is important to note that the knowledge formalized in an ontology can benefit the user community that deals with multimedia documents in a number of ways. These include compiling intelligent queries with the possibility of expanding new concepts related to the initial query; helping to formulate information needs with automated reference services, including copyright information relating to digital media; supporting automated document annotation; and improving the user experience in semantic navigation between documents displayed on the web interface as search results.

Finally, it is important to underscore that the OMRM is, above all, a recommendation based on conceptual modeling principles involving multimedia ontologies and should, therefore, be implemented, coded, tested and validated in the field to obtain consistent conclusions on its applicability. We aim at applying the OMRM in studies on the inte-

gration of databases belonging to multimedia content providers with open semantic platforms (such as Wikidata), with a view to the creative collaboration of knowledge networks. This integration would make it possible to implement and test different federated queries with useful inference mechanisms based on previously established functional requirements, with a view to obtaining more conclusive results.

Notes

1. https://mpeg.chiariglione.org/standards/mpeg-7
2. https://mpeg.chiariglione.org/standards/exploration
3. https://www.rdaregistry.info/Elements/
4. https://www.rdaregistry.info/termList/
5. https://www.rdatoolkit.org/

References

Almeida, Maurício Barcellos and Fernanda Farinelli. 2017. "Ontologies for The Representation of Electronic Medical Records: The Obstetric and Neonatal Ontology". *Journal of the Association for Information Science and Technology* 68, no. 11: 2529-42. https://doi.org/10.1002/asi.23900

Andrews, Pierre, Ilya Zaihrayeu and Juan Pane. 2012. "A Classification of Semantic Annotation Systems". *Semantic Web* 3, no. 3: 223-48. https://doi.org/10.3233/SW-2011-0056

Bekiari, Chryssoula, George Bruseker, Erin Canning, Martin Doerr, Philippe Michon, Christian-Emil Ore, Stephen Stead, Athanasios Velios, eds. 2024a. *Definition of the CIDOC Conceptual Reference Model*. Version 7.1.3. https://www.cidoc-crm.org/Version/version-7.1.3

Bekiari, Chryssoula, Martin Doerr, Patrick Le Bœuf, and Pat Riva. 2024b. *LRMoo object-oriented definition and mapping from the IFLA Library Reference Model*. https://www.cidoc-crm.org/lrmoo/ModelVersion/version-1.0-2

Bizer, Christian, Tom Heath and Tim Berners-Lee. 2009. "Linked Data - The Story So Far". *International journal on Semantic Web and information systems* 5, no. 3: 1–22. https://doi.org/10.4018/jswis.2009081901.

Borgo, Stefano and Claudio Masolo. 2009. "Foundational choices in DOLCE". In *Handbook on ontologies*, edited by S. Staab and R. Studer, 361-81. Berlin, Heidelberg: Springer Berlin Heidelberg.

Catalano, Chiara Eva, Valentina Vassallo, Sorin Hermon and Michela Spagnuolo. 2020. "Representing Quantitative Documentation of 3D Cultural Heritage Artifacts with CIDOC CRMdig". *International Journal on Digital Libraries* 21, no. 4: 359-73. https://doi.org/10.1007/s00799-020-00287-3

Charles, Valentine, Robina Clayphan, and Antoine Isaac, eds. 2017. *Definition of the Europeana Data Model v5.2.8*. https://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation//EDM_Definition_v5.2.8_102017.pdf

Dijkshoorn, Chris, Lizzy Jongma, Lora Aroyo, Jacco Van Ossenbruggen, Guus Schreiber, Wesley Ter Weele and Jan Wielemaker. 2018. "The Rijksmuseum Collection as Linked Data". *Semantic Web* 9, no. 2: 221-30. https://doi.org/doi: 10.3233/SW-170257

Europeana Tech. 2021. AI in relation to GLAMs Task Force: report and recommendations. https://pro.europeana.eu/files/Europeana_Professional/Europeana_Network/Europeana_Network_Task_Forces/Final_reports/AI%20in%20relation%20to%20GLAMs%20Task%20Force%20Report.pdf

Falbo, Ricardo de Almeida. 2014. "SABiO: Systematic Approach for Building Ontologies". In *Ontologies in Conceptual Modeling and Information Systems Engineering: Proceedings of the 1st Joint Workshop on Ontologies in Conceptual Modeling and Information Systems Engineering 21 September 2014,* edited by Giancarlo Guizzardi, Oscar Pastor, Yair Wand, Sergio de Cesare, Frederik Gailly, Mark Lycett and Chris Partridge. https://ceur-ws.org/Vol-1301/ontocomodise2014_2.pdf

Fernández-López, Mariano, Asunción Gómez-Pérez, and Natalia Juristo. 1997. "Methontology: from ontological art towards ontological engineering". In *AAAI-97 Spring Symposium Series: Proceedings of the Ontological Engineering AAAI-97 Spring Symposium Series 24-26 March 1997, Stanford, CA, USA*. 33-40. http://oa.upm.es/5484/

Ferrada, Sebastián, Nicolás Bravo, Benjamin Bustos, and Aidan Hogan. 2018. "Querying Wikimedia Images Using Wikidata facts". In *WWW '18: Companion Proceedings of the The Web Conference 23-27 April 2018 Lyon, France*, 1815-21. http://sferrada.com/pdf/wikiworkshop2018.pdf

Fink, Eleanor E. 2018. *American Art Collaborative (AAC): Linked Open Data (LOD) Initiative: Overview and Recommendations for Good Practices*. http://americanartcollaborative.org/wp-content/uploads/2018/03/AAC_LOD_Overview_Recommendations.pdf

Gangemi, Aldo and Valentina Presutti. 2009. "Ontology design patterns". In *Handbook on ontologies*, edited by S. Staab and R. Studer, 221-43. Berlin, Heidelberg: Springer Berlin Heidelberg.

Gilliland, Anne. J. 2016. "Setting the Stage". In *Introduction to Metadata,* edited by Murtha Baca. 3rd ed. Los Angeles: Getty Research Institute. http://www. getty.edu/publications/intrometadata/setting-the-stage/

Grüninger, Michael, and Mark S. Fox. 1995. *Methodology for the design and evaluation of ontologies*. http://www.eil.

utoronto.ca/wp-content/uploads/enterprise-modelling/papers/gruninger-ijcai95.pdf

Guizzardi, Giancarlo. 2020. "Ontology, Ontologies and the 'I' of FAIR". *Data Intelligence* 2, no. 1-2: 181-91. https://doi.org/10.1162/dint_a_00040

Higgins, Sarah. 2011. "Digital Curation: The Emergence of a New Discipline". *International Journal of Digital Curation* 6, no. 2: 78-88. https://doi.org/10.2218/ijdc.v6i2.191

IFLA (International Federation of Library Associations and Institutions). 2009. *Functional Requirements for Bibliographic Records: Final Report as Amended and Corrected Through February 2009*. https://repository.ifla.org/handle/20.500.14598/811

Galeffi, Agnese, María Violeta Bertolini, Robert L. Bothmann, Elena Escolano Rodríguez and Dorothy McGarry. 2016. *Statement of International Cataloguing Principles (ICP)*. Edition with minor revisions. Den Haag: IFLA. https://repository.ifla.org/handle/20.500.14598/80

Kudumakis, Panos, Thomas Wilmering, Mark Sandler, Víctor Rodríguez-Donce, Laurent Boch and Jaime Delgado. 2019. *MPEG Intellectual Property Rights Ontologies*. https://mpeg.chiariglione.org/standards/mpeg-21

Lancaster, Frederick Wilfrid. 1986. *Vocabulary control for information retrieval*. 2nd ed. Virginia: IRP.

Lemos, Daniela Lucas da Silva and Renato Rocha Souza. 2020. "Knowledge Organization Systems for the Representation of Multimedia Resources on the Web: A Comparative Analysis". *Knowledge Organization* 47, no. 4: 300-19. https://doi.org/10.5771/0943-7444-2020-4-300

Lemos, Daniela Lucas da Silva, Dalton Lopes Martins, Asla Medeiros e Sá, Luciana Conrado Martins and Danielle do Carmo. 2022. "A Proposal in Creating a Semantic Repository for Digital 3D Replicas: The Case of Modernist Sculptures in Public Spaces of Rio De Janeiro". *Knowledge Organization* 49, no. 3: 151-71. https://doi.org/10.5771/0943-7444-2022-3-151

Machado, Luis Miguel Oliveira, Renato Rocha Souza and Maria da Graça Simões. 2019. "Semantic Web or Web of Data? A Diachronic Study (1999 to 2017) of The Publications of Tim Berners-Lee and The World Wide Web Consortium". *Journal of the American Society for Information Science and Technology* 70, no. 7: 701–14. https://doi.org/10.1002/ asi.24111

Martínez, José M., Rob Koenen and Fernando Pereira. 2002. "MPEG-7: The Generic Multimedia Content Description Standard, Part 1". *IEEE multimedia* 9, no. 2: 78-87. https://doi.org/10.1109/93.998074

Martins, Dalton Lopes, Daniela Lucas da Silva Lemos, Luis Felipe Rosa de Oliveira, Joyce Siqueira, Danielle do Carmo and Vinicius Nunes Medeiros. 2022. "Information Organization and Representation in Digital Cultural Heritage in Brazil: Systematic Mapping of Information Infrastructure in Digital Collections for Data Science Applications". *Journal of the Association for Information Science and Technology* 73, no. 6: 1–20. https://doi.org/10.1002/asi.24650

Masolo, Claudio, Stefano Borgo, Aldo Gangemi, Nicola Guarino and Alessandro Oltramari. 2003. *Ontology Library (Wonder-Web Deliverable D18)*. http://www.loa.istc.cnr.it/old/Papers/D18.pdf

Messaoudi, Tommy, Philippe Véron, Gilles Halin and Livio De Luca. 2018. "An Ontological Model for the Reality-based 3D Annotation of Heritage Building Conservation State". *Journal of Cultural Heritage* 29: 100-12. https://doi.org/10.1016/j.culher.2017.05.017

Mora-Cantallops, Marçal, Salvador Sánchez-Alonso and Elena García-Barriocanal. 2019. "A Systematic Literature Review on Wikidata". *Data Technologies and Applications* 53, no. 3: 250-68. https://doi.org/10.1108/DTA-12-2018-0110

Nack, Frank, Jacco van Ossenbruggen and Lynda Hardman. 2005. "That Obscure Object of Desire: Multimedia Metadata on the Web, Part 2". *IEEE MultiMedia* 12, no.1: 54-63. https://doi.org/10.1109/MMUL.2005.12

Navarrete, Trilce, and Elena Villaespesa. 2021. "Image-based Information: Paintings in Wikipedia". *Journal of Documentation* 77, no.2: 359-80. https://doi.org/10.1108/JD-03-2020-0044

NISO (National Information Standards Organization). 2005. *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies. ANSI/NISO Z39.19-2005*. Baltimore: NISO.

Riva, Pat, Patrick LeBoeuf and Maja Žumer. 2017. *IFLA Library Reference Model: A Conceptual Model for Bibliographic Information*. https://www.ifla.org/files/assets/cataloguing/frbr-lrm/ifla-lrm-august-2017_rev201712.pdf

Robledano-Arillo, Jesús, Diego Navarro-Bonilla and Julio Cerdá-Díaz. 2020. "Application of Linked Open Data to The Coding and Dissemination of Spanish Civil War Photographic Archives." *Journal of Documentation* 76, no. 21: 67-95. https://doi.org/10.1108/JD-06-2019-0112

Shadbolt, Nigel, Tim Berners-Lee and Wendy Hall. 2006. "The Semantic Web Revisited". *IEEE intelligent systems* 21, no. 3: 96-101. https://doi.org/10.1109/MIS.2006.62

Saathoff, Carsten and Ansgar Scherp. 2010. "Unlocking the Semantics of Multimedia Presentations in the Web with the Multimedia Metadata Ontology." In *Proceedings of the 19th International Conference on World Wide Web, 26-30 April 2010 Raleigh NC, USA,* 831-40. New York: ACM. https://doi.org/10.1145/1772690.1772775

Silva, Daniela Lucas da, Renato Rocha Souza and Maurício Barcellos Almeida. 2012. "Ontologies and controlled vo-

cabulary: comparison of building methodologies". In *Semantic Technologies for Business and Information Systems Engineering: Concepts and Applications,* edited by Stefan Smolnik, Frank Teuteberg and Oliver Thomas, 1-15. Hershey, PA: IGI Global.

Siqueira, Joyce and Dalton Lopes Martins. 2021. "Workflow Models For Aggregating Cultural Heritage Data on The Web: A Systematic Literature Review". *Journal of the Association for Information Science and Technology* 73, no. 2: 204–24. https://doi.org/10.1002/asi.24498

Suárez-Figueroa, Mari Carmen, Asunción Gómez-Pérez and Mariano Fernández-López, eds. 2012. "The NeOn Methodology for Ontology Engineering." In *Ontology Engineering in a Networked World,* edited by Mari Carmen Suárez-Figueroa, Asunción Gómez-Pérez, Enrico Motta and Aldo Gangemi, 9-34. Berlin: Springer, 9-34.

Svenonius, Elaine. 2000. *The intellectual foundation of information organization*. Cambridge: The MIT Press.

van Ossenbruggen, Jacco, Frank Nack and Lynda Hardman. 2004. "That Obscure Object of Desire: Multimedia Metadata on The Web, Part-1". *IEEE MultiMedia* 11, no.4: 38-48. https://doi.org/10.1109/MMUL.2004.36

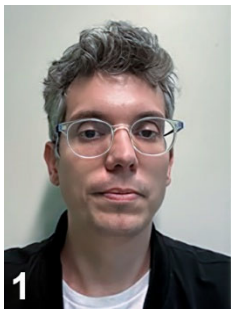
Wilkinson, Mark, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niglas Blomberg *et al*. 2016. "The FAIR Guiding Principles for scientific data management and stewardship". *Scientific data* 3, no. 1. https://doi.org/10.1038/sdata.2016.18

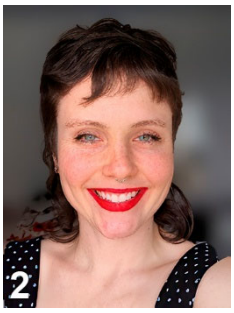
Zeng, Marcia Lei and Jian Qin. 2016. *Metadata*. 2nd ed. Atlanta: ALA Neal-Schuman.

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.

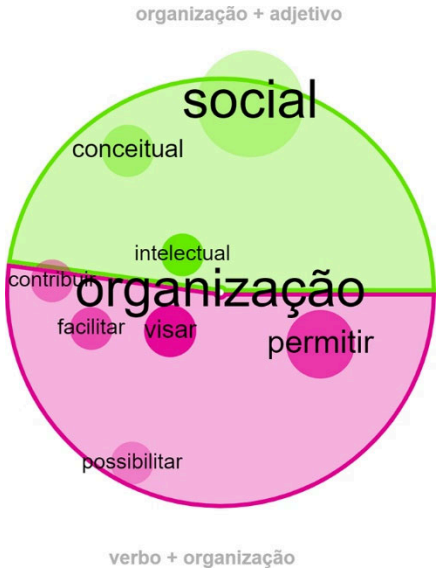


Figure 1. Word Sketch "organization".

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".

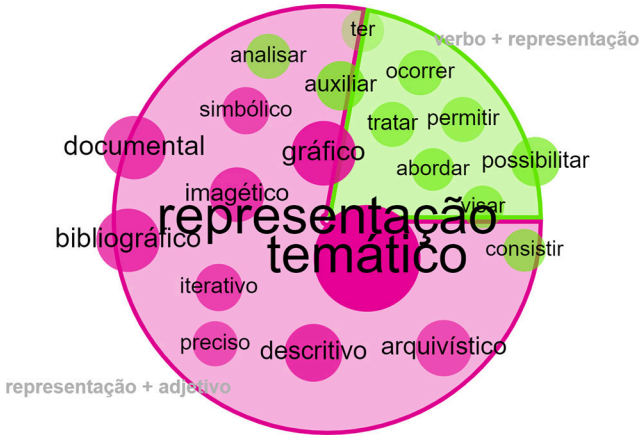


Figure 3. Word Sketch "representation".

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*.

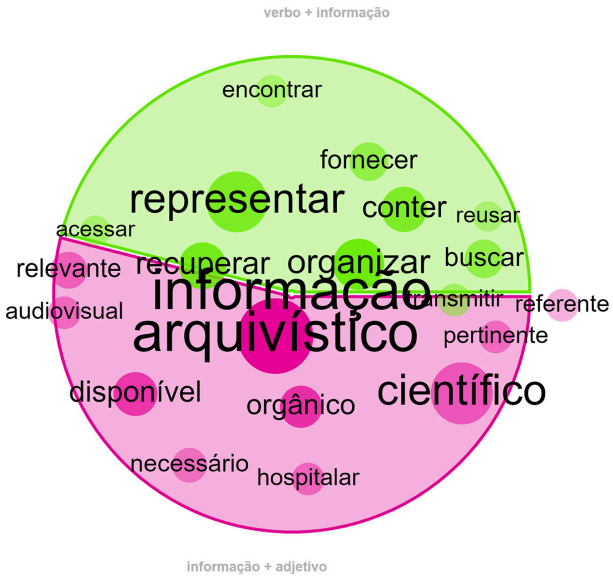


Figure 4. Word Sketch "information".

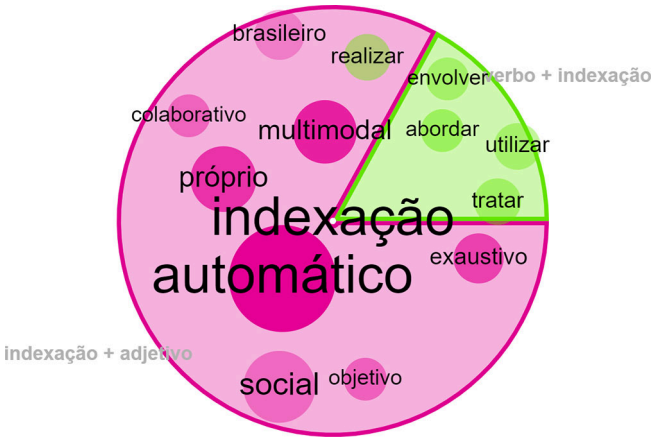


Figure 5. Word Sketch "indexing".

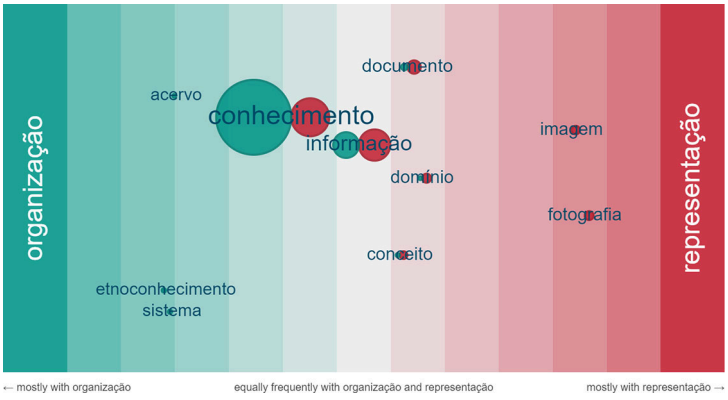


Figure 6. WSD (KNOWLEDGE-REPRESENTATION).

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.

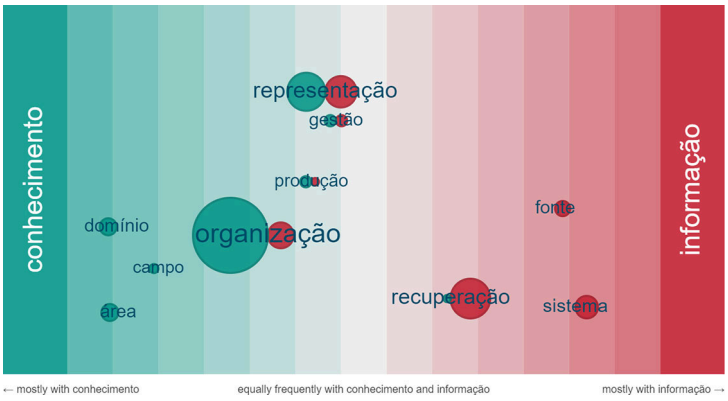


Figure 7. WSD (KNOWLEDGE-INFORMATION).

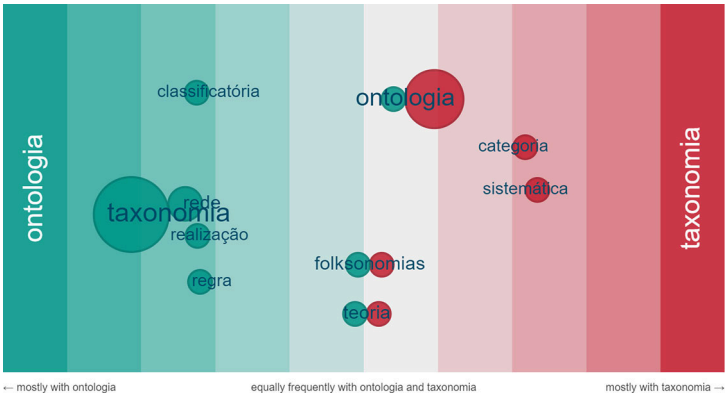


Figure 8. WSD (Ontology-Taxonomy).

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-*

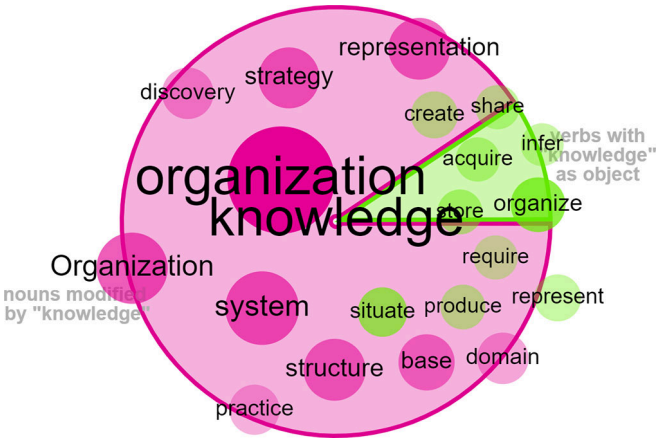


Figure 9. Word Sketch "knowledge".

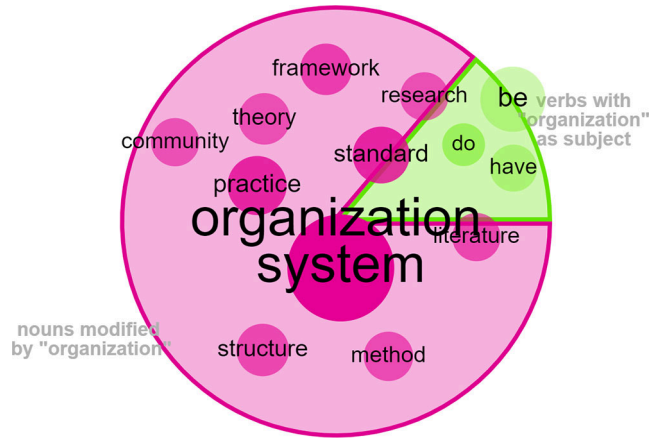


Figure 10. Word Sketch "organization".

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

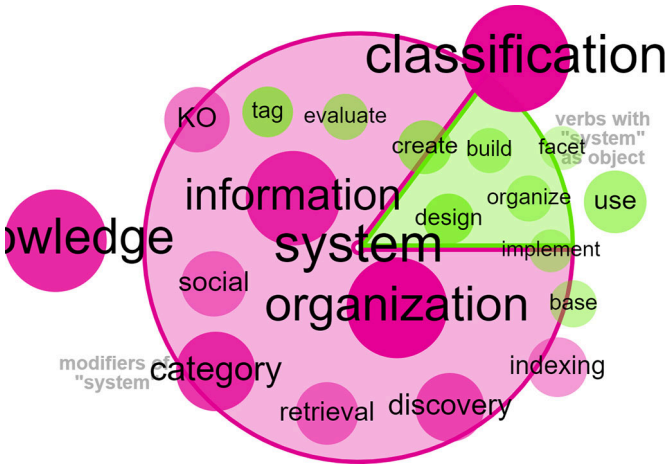


Figure 11. Word Sketch "system".

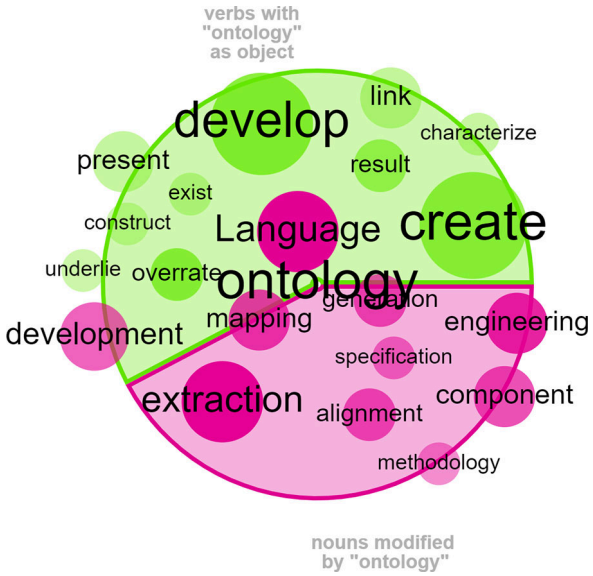


Figure 12. Word Sketch "ontology".

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.

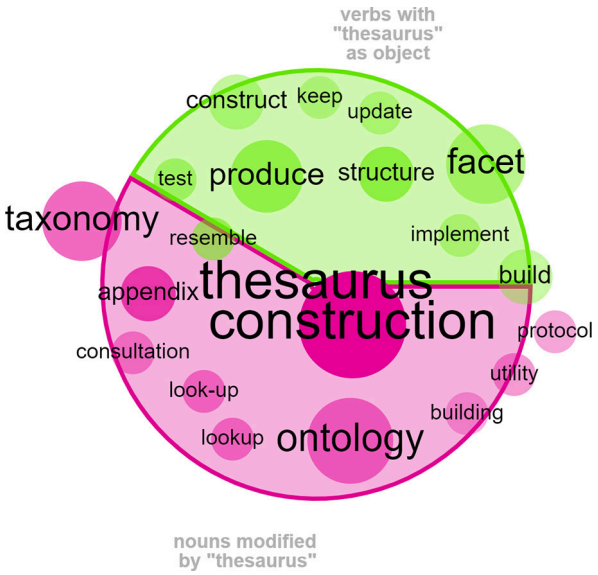


Figure 13. Word Sketch “thesaurus”.

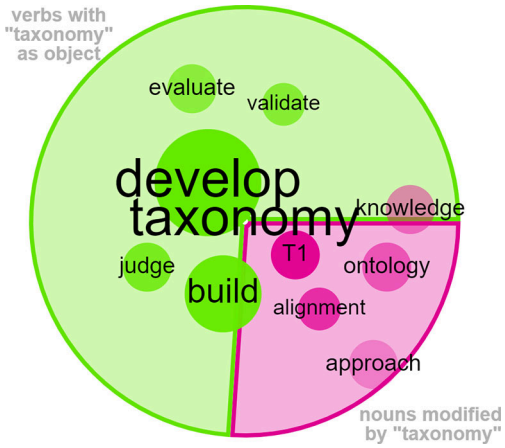


Figure 14. Word Sketch “taxonomy”.

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.

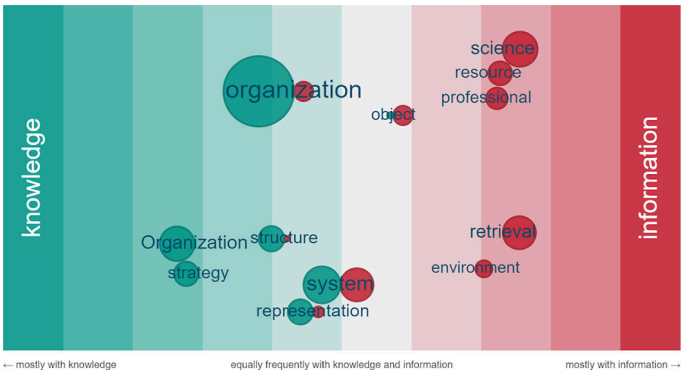


Figure 15. WSD (knowledge-information).

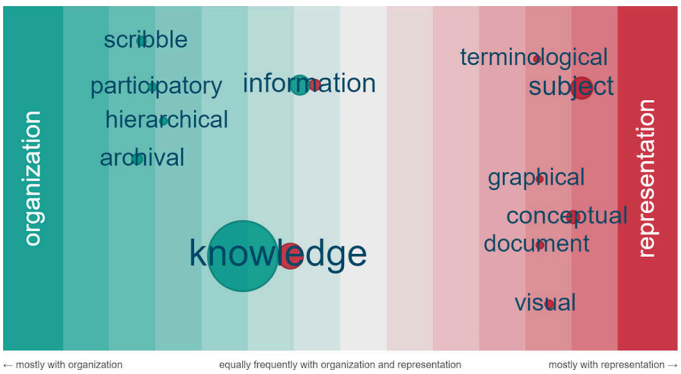


Figure 16. WSD (organization-representation).

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-

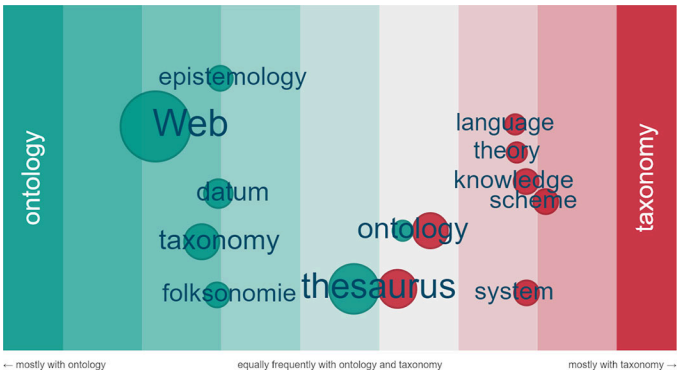


Figure 17. WSD (ontology-taxonomy).

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

Critical Archival Studies: Exploring an Emerging Domain

Anderson Berbesi* and Natália Tognoli**

* **Fluminense Federal University – Graduate Program in Information Science (PPGCI/UFF), Niterói/RJ – Brazil

*aberbesi@id.uff.br, **nataliatognoli@id.uff.br

Anderson Berbesi is a Master's student in the Graduate Program in Information Science at the Federal Fluminense University (PPGCI/UFF) and is member of the research group Archives, Libraries and Knowledge Organization, at Fluminense Federal University. He holds a bachelor's degree in Library Science from the Central University of Venezuela (UCV). His current research seeks to explore Critical archival studies as an emerging field within the archival studies and as an emergent domain in the context of domain analysis, particularly within the fields of Information Science and Knowledge Organization.

Natália Tognoli is a Professor in Department of Information Science at Fluminense Federal University, Brazil. She holds a bachelor's degree in Archival Science, a Master's, and a Ph.D. in Information Science. Her current research interests are the relationships between knowledge organization and archival science, focusing on critical studies and social justice. She coordinates the Information Science Graduation Program at Fluminense Federal University and is Editor-in-Chief of the Knowledge Organization Journal. Orcid: orcid.org/0000-0002-0609-498X

Berbesi, Anderson, and Natália Tognoli. 2024. "Critical Archival Studies: Exploring an Emerging Domain". *Knowledge Organization* 51, no. 8: 600-609. 32 references. DOI:10.5771/0943-7444-2024-8-600.

Abstract: This article explores critical archival studies as an emerging trend in the archival field, analyzing its relationship with domain analysis (DA), a methodological-theoretical approach within knowledge organization (KO). The objectives are: 1) to position critical archival studies as a research area of interest for KO, and 2) to evaluate their potential as an emergent domain according to DA. The methodology employed was a literature review, prioritizing seminal works on DA, critical archival studies, and their theoretical intersection. The findings highlight that critical archival studies, conceptualized as a theoretical-practical trend, aim to transform archival dynamics through critical perspectives but face challenges in its theoretical, institutional, and intellectual structuring, complicating its classification as a traditional domain. This prompts a discussion about characterizing critical archival studies as an emergent domain, understood as a type of domain in the phase of specialization and formation that features three main characteristics: 1) conceptual and epistemological heterogeneity, 2) developing organizational structures, and 3) consolidation potential. The study concludes that KO is a key discipline for supporting the establishment of emergent domains like critical archival studies, addressing both their external structures (e.g., groups, institutions) and internal structures (e.g., theories, discourses). This analysis broadens the discussion on emergent domains and strengthens the study of critical archival studies as a field in development.

Received: 15 September 2024; Revised 31 November 2024; Accepted 4 December 2024**.**

Keywords: critical archival studies; domain analysis; critical theories; emergent domain.

1.0 Introduction

Critical archival studies represent an emerging trend in the archival field. They spark debates and divergent positions regarding the role of archives and archivists in addressing social, cultural, political, and ethical issues. These issues often relate to the histories and memories of groups, collectives, and individuals that history – and, frequently, archivists themselves – have systematically silenced.

This critical perspective seeks to rethink archives as passive information repositories and key actors in power dynamics, identity construction, and representation. Within this framework, Critical archival studies has gained recognition in sociology, memory studies, art, history, and, more recently, Information Science (IS) (Alencar et al. 2023b; Salerno 2024). These fields share common concerns about the role of memory and the active participation of professionals in addressing the harm caused by memory institutions.

Despite growing interest, this trend has been subject to limited theoretical reflection as an entity in and of itself, both within their originating field and in related disciplines. Their conceptualization through the lens of knowledge organization (KO), mainly using domain analysis (DA) as a specific approach, remains virtually unexplored. Yet, KO and DA offer key tools to understand the internal and external dynamics of emerging trends such as critical archival studies.

In this context, exploring the relationship between critical archival studies and DA becomes essential, understanding the latter as a methodological-theoretical approach within KO.

This article is based on two main premises: 1) the need to problematize critical archival studies as a research area of interest within KO; and 2) the opportunities and challenges posed by analyzing an emerging entity such as critical archival studies through the lens of DA.

This exploration leads to a discussion of whether critical archival studies can be characterized as an emergent domain, revisiting Barité's (2020) and Tognoli (2024) ideas on this concept.

To achieve this, the article is divided into three sections:

1. Contextualizing and characterizing critical archival studies as an emerging trend within archival studies.
2. Discussing the main features and elements that define DA as a methodological approach within knowledge organization.
3. Delimiting the concept of domain by analyzing and discussing its main characteristics. This discussion addresses whether critical archival studies can be considered a domain, highlighting the difficulties in applying a traditional domain perspective and introducing the concept of emergent domains as an alternative framework.

This work employed a literature review as its primary methodology, prioritizing seminal works on domain analysis in KO, theoretical explorations of critical archival studies, and research examining the intersection of critical archival studies and domain analysis. The findings are presented narratively in the following sections.

2.0 Critical archival studies

Critical archival studies, also known as critical archival science, represent a new theoretical-practical approach established within the archival field in 2017. Its emergence was marked by a special issue of the *Journal of Critical Library and Information Studies*, introduced by Caswell et al. (2017), who framed it as an analytical stance on the social, political, cultural, and ethical dimensions of archives, particularly in relation to the dynamics of knowledge production and identity construction.

Although this idea is not entirely new, the term was first articulated in the Anglophone world in 2010, its influences can be traced back to three pivotal moments: 1) Howard Zinn's (1977) earlier speech, shaped by the radical social history movements of the 1960s and 1970s; 2) Terry Cook's (Cook 2001; Cook and Schwartz 2002) postmodern ideas from the late 20th and early 21st centuries, which expanded archival concepts and practices; and 3) Andrew Flinn's proposals for community archives since 2009 (Flinn et al. 2009), which redefined the role of communities as document producers, representing their memory and history (Caswell et al. 2017; Caswell 2021; Gustavson and Nunes 2023).

However, critical archival studies introduces added value by being the first instance in which a variety of archival perspectives – unified under a single umbrella term – seek to question and promote the transformation of archival practices within the preexisting order (Hoyle 2023). As Botnick (2019) highlights, this step was crucial for identifying a lineage of past and future studies that "interrogate, rethink, and reframe archival concepts in critical ways. The act of naming was a formal recognition of scholarship that interrogated dominant archival concepts prior to 2017 and a call to action for archivists to continue this work" (153).

In particular, the publication of the special issue of the *Journal of Critical Library and Information Studies*, as Gustavson and Nunes (2023) argue, established a framework to analyze institutional power, white supremacy within traditional archives, and strategies for promoting liberatory practices in archival stewardship. This issue strengthened "our understanding of critical conversations about whose materials archives collect, where historical oppressions are upheld and where they might be dismantled within archival institutions, and about new practices for recreating and reenvisioning the archives of the future" (7).

This trend presents a clear analytical and emancipatory purpose, capable of transforming both the archival field and society at large through a liberating praxis that opposes oppression (Caswell 2021). As Botnick (2019) explains, "CAS [Critical archival studies] is a call to action to examine power in record creation, keeping, and outreach. By breaking down what is taken for granted in this field, archivists might build a new archival practice that is liberating rather than oppressive" (153).

Following this perspective, critical archivists must advocate for a praxis that seeks to alter the dynamics of narratives and historically marginalized communities, centering their efforts on transforming the archive from a space of power and control into one of integration, cooptation, or adaptation (Alencar et al. 2023b; Hoyle 2023; Cifor et al. 2023).

This approach explicitly reflects the influence of Critical Theory ideals from the Frankfurt School, particularly Max Horkheimer's work, which inspires the field's three main ap-

proaches: theoretical, practical, and normative. However, as Caswell et al. (2017) observe, this model incorporates diverse methodologies and theoretical frameworks, including critical race theory, postmodernism, and queer theory. While these approaches extend beyond the neo-Marxist model of the Frankfurt School, they connect critical archival studies with broader epistemological currents referred to as post-critical, critical theories, or social critical theories.

In another way, since its formulation, critical archival studies have integrated diverse proposals from various fields. Beyond enriching the archival domain, it has become a lens for other disciplines that rely on archives as tools of analysis or objects of study, particularly in examining the construction of memory, culture, and history (Botnick 2019; Salerno 2024). These disciplines see in this trend not only ideas and objectives but also methodologies and tools for analyzing their realities. This demonstrates the trend's impact both within the archival field and across the broader domain of the social sciences.

Efforts to characterize critical archival studies as an autonomous field are on the rise. These efforts not only rely on external disciplines but also stem from within the field itself. In this context, Tognoli's (2024) work stands out for proposing five degrees of specialization within the domain of critical archival studies. These can be readily applied as overarching thematic frameworks in the field:

- Power dynamics
- Prejudices, biases, and exclusion
- Decoloniality/Postcoloniality
- Social justice and activism
- Professional ethics

Despite these advancements, the vast array of influences and references – not only theoretical but also epistemological, social, and political – define this new field as an amalgam of perspectives. However, this diversity presents a significant challenge that the field has yet to overcome: the lack of homogeneity in the discourses, proposals, and theories that can be categorized as critical.

This issue is better understood when considering that the critical archival field is a recently established community. It lacks institutions dedicated to teaching, researching, and promoting these initiatives. Furthermore, most researchers who identify as critical archival thinkers often do not share common concepts, ideas, theoretical-epistemological frameworks, or even references. In many cases, these efforts favor a distinctly Anglophone focus, and the perspectives among various scholars are often quite disparate, making dialogue and interaction between groups particularly challenging, despite certain existing points of connection.

The previous comment aligns with the arguments presented by Alencar et al. (2023b), who discuss the scarcity of texts addressing critical and decolonial studies in the context of Archival Science, both globally and particularly in the Brazilian region. The authors demonstrate that this body of scholarly work is still very recent, with limited literature, predominantly in English, published in journals and books that are not open-access. Additionally, there are few research centers or reference entities worldwide that align with this perspective, with University of California – Los Angeles (UCLA) in the United States being the closest example.

It is also important to consider that the conceptual and theoretical breadth of this field generates epistemological challenges that the community has yet to address. For instance: What can be considered critical in archival science? Is it anything that analyzes reality, or is this perspective too broad? Could other theoretical-paradigmatic trends or approaches – such as postmodern studies or community archives – be seen as subordinate to this trend when framed as an umbrella term, or are they of equal value, or even superior as direct influences? Can it be assumed that authors share common premises and understandings when conducting their studies and activities? Is there a particular inclination toward a specific theme, idea, or author guiding the discourse? These are just some of the many questions that remain unanswered.

These challenges emphasize the importance of treating critical archival studies not merely as a trend but as an object of study in its own right. Understanding its evolving boundaries, epistemological structures, and knowledge production requires a systematic framework. Here, the domain analysis emerges as a vital theoretical-methodological tool for addressing this complexity, offering to analyze critical archival studies both as an intellectual community and as a socio-epistemological construct, by examining its internal and external structures – its discourses, actors, institutions, and paradigms – AD provides a pathway to understand and articulate the dynamics shaping this trend as an emergent domain within the broader context of archival studies and Knowledge organization.

3.0 Domain analysis: a brief overview

Domain analysis (DA), as a concept within the context of knowledge organization (ko) and particularly information science (IS), has been explored and developed since 1995. That year, Hjørland and Albrechtsen (1995) published the seminal work *Toward a New Horizon in Information Science: Domain-Analysis*. In this study, the authors argue that the best way to understand information in the field is "to study the knowledge-domains as thought or discourse communities, which are parts of society's division of labor" (400).

Although this work helped popularize the term in the information field, its origins are not exclusive to IS. DA de-

rives from software engineering, where it was introduced by James Neighbors in the early 1980s. In this context, DA was defined as an activity aimed at identifying the common objects and operations of similar systems within a specific domain problem (Damus and Acuña 2019; Guimarães 2024).

While Hjørland and Albrechtsen's (1995) work does not represent the first introduction of the term to KO or IS[1], it is a critical milestone that established a novel methodological and theoretical perspective in the area. Their approach is characterized by a shift toward the contextual understanding of information, adopting a sociological vision that transcends the traditionally individual-centric cognitive approach (Guimarães 2014).

In this sense, DA was conceived as a paradigmatic alternative to overcome the physicalist and cognitivist approaches that dominated the early years of the discipline. These perspectives prioritized individualistic and internalist analyses, neglecting the social and cultural dimensions of information. In contrast, DA introduces a contextual perspective in which users are considered producers of information embedded in different cultures, social structures, and knowledge domains, connected through common languages and communicative practices. Thus, the meanings of information and knowledge are constructed through cultural, historical, and social processes shared among members of these communities (Romero Quesada 2013; Grácio 2020).

Under the socio-cognitive approach (Hjørland 2004), DA gains added value both as a research program and as a methodology within KO. This approach is characterized by two key aspects: analyzing the structures of knowledge domains and identifying the collective values and beliefs that shape their development and evolution.

Regarding the first aspect, Evangelista et al. (2022) argue that DA is essential for accurately understanding the composition and boundaries of a domain. This perspective focuses on characterizing work structures, ontologies, and communication patterns, or, in other words, analyzing "the circumstances under which activities occur and the constraints imposed by contemporary paradigms and research fronts" (7). In this context, DA conceptualizes knowledge as a social construct expressed through theories, paradigms, and epistemologies manifest in the activities and products of a scientific community. These expressions provide access to underlying information that helps reveal the structure and meanings of such domains (Guimarães and Tognoli 2015), offering tools to "uncover the contours of these nested and interrelated conceptual components of knowledge-producing domains" (Smiraglia 2015, 7).

The second aspect, also referred to as DA in a narrow sense, focuses on identifying the fundamental categories of a domain. This approach seeks to understand the perspectives, goals, values, and interests of the field by studying its theories, paradigms, and traditions (Evangelista et al. 2022; Hjørland 2024). This involves recognizing what is significant or meaningful in a specific field, facilitating the analysis of trends, patterns, processes, agents, and their relationships. In this way, DA becomes an invaluable tool for both information science and the scientific communities it studies (Guimarães and Tognoli 2015).

Despite its potential, DA faces a significant limitation that must be addressed before moving forward: the impossibility of conducting a completely neutral or a priori domain analysis. As Grácio (2020) points out, an adequate analysis requires "broad and deep knowledge of the theories of the studied domain, which means that domain analysis is not neutral, as it is always based on certain perspectives at the expense of others" (73). This assertion emphasizes that DA is inevitably influenced by the theoretical and methodological perspectives adopted by the analyst, necessitating the explicit justification of such choices.

This view is echoed by Damus and Acuña (2019), who argue that any analyst studying a knowledge community must first deeply understand the domain in question. This involves detailed knowledge of the social and cultural environment in which the domain develops, considering its practices and habits. Such an approach allows for identifying the essential elements of its structure and the internal and external interrelations that position it as a distinct sphere concerning others.

Although this perspective closely aligns with Grácio's (2020), it could be seen as somewhat naïve. It overlooks the diversity of viewpoints and inherent biases that any domain analysis may entail. As Kleineberg (2014 apud Hjørland 2016, 27) states:

> The knower as an agent of epistemic activity is always already embodied as a material organism and embedded in a social and cultural environment at a certain point in time and space. In other words, the prerequisites to create, represent, organize, and communicate knowledge or information are limited by preconditions which are investigated by theories of knowledge and constitute the epistemological dimension.

In line with this idea, every domain analysis is influenced by the analyst's social and theoretical interests, which in turn affect the outcomes. This highlights the importance of analysts explicitly justifying their theoretical and methodological decisions, as these shape the understanding of the reality being studied (Evangelista et al. 2022; Hjørland 2024).

With these elements in mind, the essential characteristics defining DA as a theoretical and methodological approach within KO have been identified. The next section explores one of the central concepts of this perspective: the domain and its possible interpretations. This analysis sets the stage

for addressing whether Critical archival studies can be considered a domain in theoretical and methodological terms.

4.0 Critical archival science: an emerging domain

To approach the idea of a domain, it is essential to begin with one of its most debated characteristics: its ambiguity. As Smiraglia (2012) argues, there needs to be a consensus on what constitutes a domain within the context of DA. Although efforts have been made to consolidate definitions, these vary according to interests and perspectives, encompassing disciplines, fields of knowledge, areas of expertise, and discourse communities, among other possibilities (Romero Quesada 2013). Below, the three most relevant definitions from the literature are briefly presented.

The first definition originates from the work of Hjørland and Albrechtsen (1995), who associate a domain with "thought communities or discourse communities integrated within the division of social labor" (400). From this perspective, a domain is linked to a theoretically coherent or socially institutionalized discourse or thought community. Members of such communities share a language, structure, and pattern of work cooperation, as well as common forms of communication and relevance criteria that reflect their objects of study (Evangelista et al. 2022; Hjørland 2024). This view highlights the importance of an established structure where activities are coordinated around shared goals and accepted norms, enabling the organization and legitimation of knowledge within the domain.

The second perspective comes from Jens-Erik Mai, who defines a domain as "a specialty area, a literary set, or a group of people working together within an organization" (Mai 2005, 605 quoted in Guimarães 2024, 649). According to this idea, a domain comprises a group of people working towards a specific goal. Here, the focus is on activities, collaboration, and the shared objectives that unite them. This grouping, in turn, has its own substance, defined by the institutional assumptions under which it develops (Barros and Laipelt 2021; Evangelista et al. 2022).

The third definition, one of the most frequently cited in the analyzed literature, comes from Smiraglia (2012, 114), who defines a domain as:

> a group with an ontological base that reveals an underlying teleology, a set of common hypotheses, epistemological consensus on methodological approaches, and social semantics. If, after the conduct of systematic analysis, no consensus on these points emerges, then neither intension nor extension can be defined, and the group thus does not constitute a domain.

This last definition offers greater clarity regarding the boundaries a domain must have. According to Smiraglia, a domain should exhibit an interaction between a coherent ontology, a unique epistemology defining its intellectual limits, and an effective discourse. As noted by Guimarães and Tognoli (2015), these characteristics only manifest within a "socially structured unit is formed" (563). The interaction between these elements depends on the pragmatic considerations of its members, including discourses, theoretical assumptions, and intersubjective agreements (Evangelista et al. 2022). Furthermore, Smiraglia (2012) emphasizes that if a clear consensus does not emerge after systematic analysis, the group cannot be considered a domain in the strict sense.

After presenting these three definitions, the divergent points between the various perceptions of the domain concept become evident. This not only reaffirms the plurality of definitions in the literature, as mentioned at the beginning of the section but also raises a different yet equally important question: which definition is most appropriate? Considering that each reflects specific interests and perspectives, this study opts to follow the definition proposed by Smiraglia (2012). This choice is based on three main reasons: first, Smiraglia's definition provides a precise delimitation of the structures and conceptual boundaries of a domain; second, it has been widely used as a reference by numerous authors within the field of DA, particularly in the Brazilian context (Guimarães 2014; Grácio 2020; Evangelista et al. 2022; Guimarães 2024); and finally, because this conception aligns with the specific goals of this analysis, which will be further elaborated in subsequent sections.

With this approach in mind, and before directly addressing whether critical archival studies can be considered a domain, it is necessary to define the intrinsic characteristics of a domain. This step will not only expand the understanding of the chosen perspective but will also provide a conceptual framework for future discussions.

The first element has already been addressed: every domain is a social construct, meaning its expression is shaped by the pragmatic considerations of its members (Smiraglia 2012). This process includes the generation of consensus around discourses, theoretical themes, and intersubjective agreements that delineate the knowledge within the field. Furthermore, it incorporates "forms of legitimation in formal expressions and models" (Evangelista et al. 2022, 6), thus establishing intellectual boundaries that differentiate one domain from another.

This description reflects the dual nature of domains noted by Hjørland (2024), who conceives them simultaneously as social and intellectual organizations. Complementing this idea with Smiraglia's (2015) proposals, it can be argued that every domain should be considered a group with a shared understanding of its knowledge base, marked by an underlying teleology and a shared goal that justifies the existence of the group.

In this context, the knowledge base of a domain, understood as its ontology (Smiraglia 2015), is the product of its members' activities and collective work. This base is identified through the analysis of the common vocabulary, whether in everyday speech or in their writings. From this perspective, domain analysis involves considering the interactions between the ontological, epistemological, and sociological priorities of the group. This situates the domain's work as productive, revealing its "critical role in both the evolution of knowledge and the comprehension of knowledge as a scientific entity" (7).

On the other hand, Smiraglia (2012) points out that the correlation between discourses, theoretical matters, and intersubjective agreements within a domain must be close and exhibit a high degree of conformity. But what happens when such consensus is absent? According to the author, the more dispersed its theoretical base, the less likely it is that we are dealing with a true domain. This leads to the argument that, for a group to be considered a domain, it must have a minimum agreement among its members regarding a set of shared assumptions and frameworks. If this consensus is lacking, it is difficult to speak of a domain in the proper sense.

In relation to this level of stability, Evangelista et al. (2022) make a significant remark on Hjørland's (2024) work, emphasizing that any domain analysis must start from a certain level of stability in its structures. This is based on a key premise: although the knowledge of a domain is given at the time of analysis, it is also constructed through the hermeneutic and investigative interests of the community that makes up the domain, which is simultaneously the object of study and the producer of knowledge. As the authors state: "These factors also determine the continuous character of science: even if one of the two elements – the institution or the content – changes during the studies, the other aspect remains stable." (5).

This idea introduces an essential characteristic of domains that deserves detailed analysis: their constant evolution. Several authors agree that domains are dynamic entities, subject to continuous change and transformation (Smiraglia 2012; Albrechtsen 2015; Evangelista et al. Guimarães 2022; Guimarães 2024; Hjørland 2024;). Tennis (2012) synthesizes this idea by emphasizing the need to operationalize domains to adapt to their changing nature:

> What we hope to make clear in this text is that we must understand the limits of our own analyses. The reason why this is important is linked, intimately, with the fact that everything changes. Our view of the domain changes and the domains themselves change (11).

According to Hjørland (2024), domains are neither unequivocal nor static entities; rather, they are characterized by processuality, fragmentation, and indeterminacy. Along these lines, Smiraglia (2012) argues that domains are dynamic because they play a symbiotic role in the evolution of both knowledge spaces and the real world. This idea can be expanded further by recognizing that domains are not photographic retentive systems, they are never frozen "in time and space but are always changing, even if it does not appear so to producers, users, or information mediators in everyday academic practice" (Hjørland 2024).

With these elements in mind, domains, at least from Smiraglia's perspective (2012; 2015), can be understood as social constructions with an internal teleology founded on four essential pillars: the ontological, epistemological, methodological, and semantic. However, although domains are often presented as relatively stable entities due to their organizational and intellectual structures, these characteristics are not immutable. Domains are in constant flux, driven by contextual dynamics and the interests of their communities.

This leads to a key question: Can critical archival studies be considered a fully realized domain?

To answer this question, it is worth noting that efforts have already been made to relate critical archival studies to DA. Among these are the works of Alencar et al. (2023b) and Tognoli (2024). The former seeks to explore the international editorial domain of archival studies, identifying critically and decolonially oriented works that reveal an emerging scientific and epistemic-bibliographic structure within the international archival field. Meanwhile, Tognoli (2024) identifies Critical archival studies as an emerging domain, deriving five degrees of specialization from a historical and epistemological analysis that could, in the future, "comprehend the domain's configuration through the study of its epistemic and discursive communities" (955).

Both studies are prospective and aim to establish a foundation for understanding and discussing the structure of Critical archival studies as a domain. However, Tognoli (2024) also warns that it is a community still under construction, both theoretically and organizationally. This suggests that this trend lacks a stable base, whether institutional or intellectual, which allows us to revisit the initial question.

If we assume that every domain must have a coherent and stable structure, grounded in a set of pillars derived from intersubjective consensus, Critical archival studies can hardly be considered a domain in the traditional sense. This is because the fundamental elements that define a domain do not align with the intrinsic characteristics of them.

This idea becomes clearer when analyzing the essential aspects of any domain, starting with the most basic: its name and its definition. As Tennis (2012) notes, any operationalization of a domain worthy of analysis requires basic elements for its identification, the first of which is its name.

In the case of Critical archival studies, the name presents no significant difficulties, as the possible semantic ramifi-

cations or derivations maintain a common link through the particle *critical*. This semantic connection is reflected in the various denominations adopted globally, such as *critical archival studies*, *critical archival science*, or *critical archiving and recordkeeping*. This fact evidences a certain level of agreement within the community, at least in this aspect.

The second element, the definition, is considerably more problematic. On one hand, this can be attributed to the limited theoretical efforts aimed at understanding what Critical archival studies are; on the other, to the multiplicity of perceptions within the community. Among American authors, some understand it as a trend born from an intellectual evolution since 1977; others consider it a subfield within Archival Studies; still others interpret it as a theory, while in Australia, under the denomination *critical archiving and recordkeeping*, it is perceived as a methodology based on Records Continuum Theory.

The conceptual divergence surrounding Critical archival studies not only complicates the attainment of a common definition to unify the community under a shared framework but also hinders the possibility of establishing agreements among its members. This issue highlights the limitations of considering it a domain in the traditional sense. However, for analytical purposes, a provisional definition is proposed to advance this discussion. In this study, Critical archival studies is understood as a theoretical-practical trend within the archival field that adopts an analytical stance on the social, political, cultural, and ethical dimensions of archives, particularly in their relationship with knowledge production and identity construction dynamics. This approach is marked by a diversity of theories, epistemological positions, and methodologies that, although often contradictory, converge in a common goal: to question the dynamics of domination and power exerted by and through archives, with the aim of promoting social emancipation and transforming archival structures.

On this basis, a second key question arises: Is there a clear ontological, epistemological, methodological, and semantic foundation in Critical archival studies that would allow its consolidation as a domain? To answer this, the essential questions posed by Smiraglia (2015) about domains are taken as reference.

Does the group share a common goal that is implicit or explicit in its knowledge base (ontology)? Defining a clear ontological foundation within the critical archival community is challenging. While shared objectives can be identified – questioning, emancipation, and transformation within and for archives – the ways these objectives are addressed often differ. Members of the community adopt specific aspects of these goals and analyze them from perspectives that do not always allow for consensus. For instance, criticism of archives as tools does not necessarily lead to a search for transformation or emancipation, and vice versa. This significant fragmentation raises an interesting question: Is this flexibility a strength that enables the inclusion of multiple perspectives, or a limitation that hinders the domain's consolidation?

Is there a theoretical paradigm in operation that unifies a set of shared hypotheses (epistemology)? A defining characteristic of critical archival studies is its epistemological fluidity. While some works have identified minimal epistemological elements to conceptualize this trend – such as the proposals of Frankfurt School Critical Theory (Caswell et al. 2017) or social critical theories – these encompass diverse approaches. These approaches often include perspectives that, while converging in their analysis of society, are divergent and frequently contradictory, as seen in the contrast between Habermasian thought and Lyotardian postmodernism. Therefore, critical archival studies cannot be said to have an exclusive theoretical paradigm but rather a plurality of approaches that may share related hypotheses without being unified.

Is there a methodological consensus? The epistemological breadth within the critical archival community translates into considerable methodological diversity. In fact, foundational texts have promoted this diversity as an essential characteristic of critical archival studies. Nevertheless, there appears to be an implicit consensus favoring qualitative methodologies focused on understanding rather than explanation. This methodological focus, while not exclusive, seems to align with the critical and transformative nature that defines this trend.

Do critical archival studies share a social semantics, understood as a set of terms and meanings shared by its community? It is evident that certain terms, such as decolonization, representation, and power in archives, are common within the trend. However, their definitions are not always agreed upon. Given the relatively recent emergence of this trend, its social semantics is still under construction, consolidating through semantic patterns disseminated in both formal and informal academic contexts.

This point brings us to another dimension of analysis: Does an institutional organization exist around Critical archival studies? Studies by Alencar et al. (2023a) and Alencar et al. (2023b) have explored this question in two specific contexts: the editorial, analyzing journals and series specializing in critical topics, and the institutional, examining the role of the Archival Education and Research Institute (AERI) as an entity that includes specialists in critical studies among its members. Despite these efforts, there is no clear institutional organization solely dedicated to advancing, researching, and evolving this trend. Additionally, no epistemic communities or influential researcher networks have been identified in this field thus far.

Given these elements, can critical archival studies be considered a domain? Following Smiraglia's ideas (2012; 2015),

the lack of theoretical and institutional structure and cohesion, combined with its evolving ontological, epistemological, methodological, and semantic dynamism and the limited efforts of its community to define its boundaries and intrinsic characteristics, suggests that critical archival studies cannot yet be regarded as a domain in the classical sense.

However, this conclusion does not diminish the relevance of it. Instead, it highlights the need to analyze it as a domain under construction, open to defining its own boundaries and structures. This requires considering new perspectives to address it through domain analysis. In this context, Tognoli (2024) offers a promising conceptualization, proposing critical archival studies as an emergent domain, noting that such domains "develop from the intersection of disciplines like archival studies, cultural criticism, social and political studies, history, knowledge organization, and others, in response to social demands" (954).

The notion of emergent domains is not unique to Tognoli but follows Barité's (2020) ideas, who defines emergent domains as a type of domain developed in recent times (20 to 40 years) "as a result of the rapid process of specialization and reciprocal intersection between disciplines and/or thematic fields" (245). According to Barité, in their early decades, such specializations are "in the process of shaping and specifying" (245). Emergent domains reflect the expansion of interdisciplinary studies, technological developments, and the evolution of intellectual thought focused on social and cultural issues arising from the post-war era.

Barité's definition of emergent domains, particularly the idea that any recently developed domain is in a phase of formation and specialization, is crucial to understanding critical archival studies. However, certain aspects of this definition do not fully align with this trend and raise several questions.

First, the temporal framework proposed (20 to 40 years) seems arbitrary, especially when considering fields that have emerged more recently, like Critical archival studies, whose explicit development began in 2017. This prompts reflection on whether the proposed timeframe is suitable for all disciplines or requires adjustments according to specific contexts.

Second, Barité (2020) links the emergence of domains to historical processes derived from the post-war period. Here, one might ask: Is this period the only one capable of explaining the emergence of domains? Would it not be more useful to think of emergent domains as phenomena responding to specific sociocultural contexts, rather than exclusively tying them to a particular historical moment?

Lastly, the author does not address whether epistemological or ontological consensus is necessary for a field to be considered a domain. If such a consensus is required, can heterogeneous fields, trends, or perspectives lacking clear cohesion be considered emergent domains? This leads to a second question: What level of consensus is necessary to classify something as an emergent domain? This question becomes even more relevant considering that, at least in our area of interest, perspectives and approaches are highly varied and have only just begun to be mapped.

The proposal to analyze critical archival studies as an emerging domain invites a reconsideration of traditional approaches to DA and creates opportunities to understand and characterize fields still in the process of formation. Drawing from the ideas of Barité (2020) and Tognoli (2024), in this work emerging domains are understood as those still undergoing processes of shaping and specification, influenced by socio-cultural contexts closely tied to the present, and defined by the following characteristics:

1. A conceptual and epistemological heterogeneity marked by ongoing debates and negotiations, where multiple perspectives coexist without achieving complete consensus.
2. An absence of clearly defined organizational structures, though with emerging indications of both institutional and intellectual organization.
3. A potential for growth and consolidation as more researchers, institutions, and knowledge outputs contribute to its development.

Once an intersubjective agreement within its community regarding topics, perceptions, institutions, methodologies, objects, among other characteristics, is identified through analysis, it will be possible to properly speak of a domain.

Following this definition, DA should not merely be perceived as a methodology aimed at breaking down existing domains but should take on a more active and constructive role in relation to emerging domains. In this context, DA must focus on understanding, delimiting, and discovering the epistemological configurations of these developing domains. This includes analyzing both the external structure (social processes, organizational structures, epistemic communities, among others) and the internal structures (predominant themes, influential theories, developing discourses) that would form their identity.

Following Albrechtsen's (2015) words, DA, in relation to emerging domains, should not be limited to describing reality, but should be actively involved in its creation: "Domains are not terrains out there, waiting to be described and analyzed by the initiated few. Fundamentally, we may all create them" (561). This implies that DA does not merely observe and record but also, at least in the case of critical archival studies, would map its potential areas of consensus and fragmentation, understand its consolidation processes, and ultimately contribute to its characterization as a domain in formation.

5.0 Conclusion

Critical archival studies, understood as a complex entity still under construction, positions itself as one of the most disruptive efforts in recent years to question and transform the traditional dynamics imposed by archives. Its connection with established fields like KO broadens the scope of its proposals, allowing them to be analyzed from perspectives that enrich and diversify existing academic discussions.

This paper argued that critical archival studies cannot be perceived as a fully consolidated domain, which requires a rethinking of the boundaries of domain analysis to address entities still in formation. In this sense, the category of emerging domain is proposed as a reference framework that allows for analyzing these still-developing trends, considering both their external structure (groups, institutions, publications) and their internal structure (discourses, theories, themes).

However, this analysis is just a starting point. Although critical archival studies has been taken as a case study, further exploration is needed to determine whether the emerging domain proposal applies to other entities with similar characteristics. Additionally, there is a need to reflect on the role of DA in the construction of emerging domains: which theoretical and methodological approaches from Hjørland (2002) are most suitable for structuring both the external and internal dimensions of these domains? Is it possible to combine different perspectives to achieve a more comprehensive and robust analysis?

These questions open up space for future research that not only expands the discussion on emerging domains but also deepens the study of critical archival studies as a developing field. KO, in this context, positions itself as a key discipline to support and foster the progress of these domains, contributing both to their analysis and establishment.

Acknowledgments

We thank the Coordination for the Improvement of Higher Education Personnel (Capes) and the National Council for Scientific and Technological Development (CNPq) for the scholarship and grant conceived for this research. We also appreciate all the careful comments given by the two anonymous reviewers and by Birger Hjørland and Thiago Barros.

Endnote

1. According to Hjørland (2024), the earliest connection can be traced back to the work of Prieto-Díaz, from the field of Software Engineering. Building on Ranganathan's faceted classification theory, Prieto-Díaz developed a synthetic analytical approach for classifying software components, which he termed domain analysis (Albrechtsen 2015).

References

Albrechtsen, Hanne. 2015. "This Is Not Domain Analysis". *Knowledge Organization* 42, no. 8: 557–61. https://doi.org/10.5771/0943-7444-2015-8-557

Alencar, Maíra Fernandes, María Brigida Nogueira Cervantes and Mario Barité. 2023a. "O contexto de uso do termo arquivos comunitários a partir da vertente dos estudos críticos e decoloniais: uma análise terminológica pontual". *Páginas a&b*, no. 19: 55–77. https://doi.org/10.21747/21836671/pag19a4.

Alencar, Maíra Fernandes, Natália Bolfarini Tognoli and Brígida Maria Nogueira Cervantes. 2023b. "Outros horizontes críticos e decoloniais nos estudos arquivísticos: em construção um mapeamento do domínio editorial internacional". *Encontros Bibli* 28: 1–26. https://doi.org/10.5007/1518-2924.2023.e92661.

Barité, Mario. 2020. "Terminología de urgencia y garantías para la representación temática: elementos para el análisis de dominios de emergencia súbita (DES)." *Informação & Informação* 25, no. 3: 239–65. https://doi.org/10.5433/1981-8920.2020v25n3p239.

Barros, Thiago Henrique Bragato and Rita do Carmo Ferreira Laipelt. 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

Botnick, Julie. 2019. "Archivists as Amici Curiae: activating Critical Archival Theory to confront racialized surveillance". *Journal of Radical Librarianship* 5, 153-72. https://journal.radicallibrarianship.org/index.php/journal/article/view/35

Caswell, Michelle, Ricardo Punzalan and T-Kay Sangwand. 2017. "Critical archival studies: An Introduction". *Journal of Critical Library and Information Studies* 1, no. 2: 1–8. https://doi.org/10.24242/jclis.v1i2.50

Caswell, Michelle. 2021. *Urgent Archives: Enacting Liberatory Memory Work*. London: Routledge.

Cifor, Marika, Carolina Alves and Martina Spohr. 2023. "Interview with Marika Cifor". *Estudos históricos* 36, no. 79: 129-40. https://periodicos.fgv.br/reh/article/download/89358/83894

Cook, Terry and Joan M. Schwartz. 2002. "Archives, records, and power: From (postmodern) theory to (archival) performance". *Archival Science* 2: 171–85. https://doi.org/10.1007/BF02435620

Cook, Terry. 2001. "Archival science and postmodernism: new formulations for old concepts". *Archival Science* 1: 3–24. https://doi.org/10.1007/BF02435636

Damus, María Arminda, and Gabriela Noemí Acuña. 2019. "Aproximación al Análisis de Dominio (AD) desde la investigación En Bibliotecología y Ciencia de la Infor-

mación". *E-Ciencias de La Información* 9 (2). https://doi.org/10.15517/eci.v9i2.37497

Evangelista, Isadora Victorino, Maria Cláudia Cabrini Grácio and José Augusto Chaves Guimarães. 2022. "The Concepts of Domain, Discourse Community and Epistemic Community: Affinities and Specificities". *Brazilian Journal of Information Science: Research Trends* 16. https://doi.org/10.36311/1981-1640.2022.v16.e02138.

Flinn, Andrew, Mary Stevens and Elizabeth Shepherd. 2009. "Whose memories, whose archives? Independent community archives, autonomy and the mainstream". *Archival Science* 9, nos. 1-2: 71–86. https://doi.org/10.1007/s10502-009-9105-2

Grácio, Maria Cláudia Cabrini. 2020. *Análises relacionais de citação para a identificação de domínios científicos: uma aplicação no campo dos estudos métricos da informação no Brasil*. Marília: Oficina Universitária.

Guimarães, José Augusto Chaves and Natália Bolfarini Tognoli. 2015. "Provenance as a Domain Analysis Approach in Archival Knowledge Organization". *Knowledge Organization* 42, no. 8: 562–69. https://doi.org/10.5771/0943-7444-2015-8-562

Guimarães, José Augusto Chaves. 2014. "Análise de Domínio como perspectiva metodológica em Organização da Informação". *Ciência Da Informação* 43, no. 1: 13–21. https://doi.org/10.18225/ci.inf.v43i1.1415

Guimarães, José Augusto Chaves. 2024. "A análise de domínio como perspectiva teórico-metodológica uma análise dos aspectos teóricos na literatura internacional na organização do conhecimento: uma análise dos aspectos teóricos na literatura internacional". *Revista Ibero-Americana de Ciência Da Informação* 17, no. 3: 644–77. https://doi.org/10.26512/rici.v17.n3.2024.53674

Gustavson, Andi and Nunes, Charlotte. 2023. "Introduction". In *Tranforming the Authority of the Archive: Undergraduate Pedagog and Critical Digital Archives,* edited by Andi Gustavson and Charlotte Nunes, 1–20. [Mountain View]: Lever Press.

Hjørland, Birger, and Hanne Albrechtsen. 1995. "Toward a New Horizon in Information Science: Domain-Analysis". *Journal of the American Society for Information Science* 46 (6): 400–25. 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. 2004. "Domain Analysis: A Socio-Cognitive Orientation for Information Science Research". *Bulletin of the American Society for Information Science and Technology* 30, no. 3: 17–21. https://doi.org/10.1002/bult.312

Hjørland, Birger. 2016. "Informetrics Needs a Foundation in the Theory of Science." In *Theories of Informetrics and Scholarly Communication*, edited by C. R. Sugimoto, 20–46. Berlin: Walter de Gruyter.

Hjørland, Birger. 2024. "Análise de Domínio". *Em Questão* 30. https://doi.org/10.1590/1808-5245.30.140568

Hoyle, Victoria. 2023. *The Remaking of Archival Values*. London: Routledge.

Romero Quesada, M. 2013. "El Análisis de dominio y sus presupuestos teóricos-metodológicos". *Bibliotecas. Anales de Investigación,* nos. 8–9: 228–38. http://revistas.bnjm.sld.cu/index.php/BAI/article/view/288

Salerno, Daniele. 2024. "An Instituting Archive for Memory Activism: The Archivo de La Memoria Trans de Argentina". *Memory Studies* 17, no. 2: 332–48. https://doi.org/10.1177/17506980221150900

Smiraglia, Richard P. 2012. "Epistemology of Domain Analysis". In *Cultural Frames of Knowledge*, edited by Richard P. Smiraglia and H. Lee, 111–24. Wurzburg: Ergon.

Smiraglia, Richard P. 2015. *Domain Analysis for Knowledge Organization: Tools for Ontology Extraction*. Oxford: Chandos Publishing.

Tennis, Joseph T. 2012. "What Does a Domain Analysis Look like in Form, Function, and Genre?". *Brazilian Journal of Information Science* 6, no. 1: 3–14. https://doi.org/10.36311/1981-1640.2012.v6n1.02.p3

Tognoli, Natália. 2024. "Estudos arquivísticos críticos: caracterização inicial de um domínio emergente na práxis arquivística." In *Diálogos Na Ciência Da Informação: Atas Do XIV Encontro EDICIC*, edited by Carlos Silva, Jorge Revez and Luís Corujo, 951–56. Lisboa: Universidade de Lisboa.

Zinn, Howard. 1977. "Secrecy, Archives and the Public Interest". *Midwestern Archivist* 2: 14-26.

Introduction to the Special Issue: Critical and Social Knowledge Organization

Suellen Oliveira Milani*, Widad Mustafa El Hadi**, and Natália Tognoli***

* *** Fluminense Federal University, Information Science Department, Campus do Gragoatá.
Bloco J, Rua Professor Marcos Waldemar de Freitas Reis, s/n - São Domingos – Niterói/RJ, Brazil
**iSchool & Geriico, University of Lille (France), University of Lille, BP 60149,
59653 Villeneuve d'Ascq Cedex, France
* suellenmilani@id.uff.br, **widad.mustafa@univ-lille.fr, *** nataliatognoli@id.uff.br

Suellen Oliveira Milani is an Associate Professor in Department of Information Science and Graduate Program of Information Science at Fluminense Federal University, Brazil. She holds a Doctoral Degree in Information Science. Among her current research interests are knowledge organization and representation, ethics in knowledge organization, theoretical foundations of information science, information user studies. Complete résumé available at lattes.cnpq.br/2821128318835622. Orcid: orcid.org/0000-0001-7183-6030

Widad Mustafa El Hadi, is Professor of Information and Communication Sciences at the University of Lille SHS where she is in charge of International Relations and of the Master Diploma program. She holds a PhD in linguistics from the University of Lyon 2, France and a Habilitation from the University of Lille in 2004. Her main areas of interest are: theoretical approaches to knowledge organization; language & culture and their impact on knowledge organization; knowledge organization systems and their evaluation; cross-language and cross-cultural information retrieval, and, more recently, Digital Humanities and ethics in information & knowledge organization. She has been active in ISKO (International Society for Knowledge Organization since 1996, as co-founder of the French ISKO Chapter (with Jacques Maniez) and elected as the first ISKO-France president (1996-2001, 2015-2023). She publishes regularly on knowledge organization, digital cultures, digital humanities and recently the Ethics of KO. She has supervised 18 doctoral theses on terminology; library and information science; information literacy; archiving and knowledge organization systems and Digital humanities. She founded in June 2021 the "International Thematic Network on Ethics in SHS".

Natália Tognoli is a Professor in Department of Information Science at Fluminense Federal University, Brazil. She holds a bachelor's degree in Archival Science, a Master's, and a Ph.D. in Information Science. Her current research interests are the relationships between knowledge organization and archival science, focusing on critical studies and social justice. She coordinates the Information Science Graduation Program at Fluminense Federal University and is Editor-in-Chief of the Knowledge Organization Journal. Orcid: orcid.org/0000-0002-0609-498X

Milani, Suellen Oliveira, Widad Mustaf El Hadi, and NatáliaTognoli. 2024. "Introduction to the Special Issue: Critical and Social Knowledge Organization". *Knowledge Organization* 51, no. 8: 610-612. 2 references. DOI:10.5771/0943-7444-2024-8-610.

Information institutions and systems have the goal to ensure access to information and records, support research, preserve and disseminate memory, promote accountability and encourage reading and cultural activities. For that, professionals construct, maintain and apply knowledge organization systems in a given context, space and time, taking into account the diversity of knowledge domains, users' needs and reflecting on their own agency.

It is a global concern of KO to create alternatives to promote spaces for the appropriate representation of knowledge domains and user communities' needs, mainly the ones that do not belong to the dominant groups, but solutions are local and/or contextualized, generally. For that, it is important to consider cultural and linguistic barriers, theoretical assumptions, values and critical and social studies. These studies comprise a topic of KO, which is suitable for hosting such discussions.

Social studies, including cultural perspectives, reflect specific discursive communities and knowledge domains. As Ibekwe (2024) observes, critical studies in KO investigate the power dynamics, biases, and social impacts embedded in the systems and methods used to classify and organize information.

Approaching a critical and social KO requires the recognition of a framework of intertwined injustices that are deeply rooted in historical practices of exclusion and silencing. These injustices support not only a social exclusion but also an epistemic one, especially if we aim to address social justice and epistemic justice in KO systems, for instance.

Social justice in KO refers to a more inclusive knowledge organization system as part of institutions that promote access to information as an essential right to support responsive, participative and representative decision-making. It means the elimination of institutionalized domination and oppression (Young 1990). Creating KO systems under a social justice perspective means doing it in a defensible, participatory and transparent way, involving historically excluded groups to avoid reinforcing institutional biases, challenging the illusion of neutrality, recognizing their ideological nature and actively promoting the preservation and visibility of diverse cultural narratives and epistemologies.

Thus, KO deals with the inherent tension between the non-neutrality of representation of knowledge and the commitment to respond promptly and accurately to the demands and information needs of different knowledge domains and user communities. This dynamic has led the field of KO to engage in a discussion regarding the promotion of access, ethics, and social justice through studies and actions. This special issue explores the assumptions underlying critical and social approaches to KO. Critical studies within KO focus on examining the field's presumptions and proposing solutions to its challenges.

From this perspective, critical and social KO enables diverse approaches. For this Special Issue, the following main topics were outlined:

a. Antiracism, anticolonialism, and feminist stances;
b. Epistemicide and epistemic (in)justice in KO systems;
c. Inclusive terminology/metadata;
d. Indigenous knowledge organization;
e. KO support for social justice;
f. Truth and relevance in KO.

The selection process for this special issue was highly competitive, with 47 relevant proposals submitted. After careful evaluation, we are honored to present 12 papers, divided across two issues (vol. 51, issues 7 and 8).

Part I focuses on racism, intersectionality, decoloniality, epistemic injustices, and marginalized communities in KO systems.

Chris Holstrom, in "Critical control: how different forms of vocabulary control aid and hinder novice indexers aiming to support racial justice", investigates how novice indexers assign terms to documents addressing race and racism. The study analyzes the faithfulness, exhaustivity, and specificity of the indexing terms while also examining how confident and in-control the indexers feel throughout the process.

Maria Aparecida Moura, in "Information and code biases: social differentiation, intersectionality, and decoloniality in KO systems", introduces an experimental framework that integrates intersectionality and decoloniality as essential dimensions in designing and structuring knowledge organization systems.

In the paper "Toward an etiology of harm for knowledge organization: onto-epistemic injustice in classificatory systems of record", Beth Patin, Tyler Youngman and Elliott Hauser discuss two instances of actual epistemic harms from knowledge organization systems tracing the effects they have upon what can be known through them, the identities they create or deny, and the resulting structure of reality they uphold.

In "Psychiatric classifications: an epistemic justice issue", Laurence Favier and Stéphanie Derdar examine the portrayal of schizophrenia within medical and bibliographic classifications, as well as French-language media. The paper highlights efforts to destigmatize schizophrenia through the evolution of classification practices.

Nick Ubels, Lisabelle Tan, and Angel Long, in "Developing person-centered metadata: a case study of the Behaviours in Dementia Toolkit" describe the process and outcomes of navigating the complexity of developing, refining, and implementing person-centered metadata in the context of the Behaviours in Dementia Toolkit.

Heather Moulaison-Sandy, Karen Snow, and Brian Dobreski, in "Social dimensions of culture, code-switching, and controlled vocabularies", delve into the concept of code-switching as a cultural phenomenon. They explore how marginalized users navigate linguistic and cultural differences in KO systems, shedding light on the challenges of creating inclusive vocabularies.

Part II addresses topics such as autism, indigenous knowledge and language, feminicide, folksonomies, and KO processes.

In the paper "Sort of people: considerations about the ontogeny of autism in the Dewey Decimal System, 1942-2023", Dóra Pákozdi traces the ontogeny of autism within the Dewey Decimal System, from the 14th to the 23rd editions highlighting the shifts in the psychiatric understand-

ing and societal attitudes toward autism, the dynamics of neurodiversity recognition, the conceptualization of autism as a spectrum, and the influential role of autistic self-advocacy.

To address a discussion regarding the subject representation of the different aspects of indigenous peoples' cultures in Latin America, Mario Barité and Mirtha Rauch present the paper "Native peoples and knowledge organization: perspective from the indigenous warrant". The study compares approaches from Canada, the United States, Australia, and Latin America to highlight best practices and challenges.

In the paper "Organization and representation of indigenous scientific production: a case study on the institutional repository in Brazil", the authors Caroline Periotto, Felipe Arakaki, Jair de Jesus Massa, Luciana de Souza Gracioso and Luzia Sigoli Fernandes Costa present a diagnosis, propositions and pre-tests in an institutional repository for the inclusion of metadata for indigenous authorship, summary and keywords in indigenous language.

Julia Bullard presents "Provocations of process in critical knowledge organization work", which discusses the following assumptions: a. systems should be seamless and not reveal the work behind them; b. systems should achieve a single authorial voice through consistency, precedent, and patterns; and c. knowledge organization systems are best applied with minimal interpretation on the part of the worker. The author argues that the most provocative work in critical knowledge organization is happening at the level of process.

In the paper "Knowledge organization systems typifying crimes of violence against women, homicide of women and femicide: a proposal", Rochelle Martins Alvorcem, Gercina Ângela de Lima and Maria Cristina Vieira de Freitas offer a guideline to typify crimes of violence against women in filling out the National Risk Assessment Form and police report through knowledge organization systems.

Finally, Sarah Barriage, Beth Strickland Bloch, and Vanessa Kitzie, in "Drag storytimes and bibliographic invisibility: a comparative analysis of picture book subject metadata", analyze the metadata associated with picture books featured in drag storytimes. The study compares metadata from library catalogs with that from online social cataloging platforms to highlight gaps and opportunities for inclusivity.

Critical and social studies on KO, including knowledge organization systems, in archives, libraries, and information institutions and environments must be constantly conducted. In this special issue, the readers will find theories, methodologies, toolkits and case studies carefully presented by specialists who have been studying KO processes, instruments, and products from a critical stance focusing on the social dimension.

References

Ibekwe, Fidelia. 2024. "Another Look at a Knowledge Organization Pioneer: Traces of Racism in Paul Otlet's Writings". *Knowledge Organization* 51, no. 1:3-18. doi.org/10.5771/0943-7444-2024-1

Young, Iris M. 1990. *Justice and the Politics of Difference*. Princeton: Princeton University Press.

Sort of People: Considerations About the Ontogeny of Autism in the Dewey Decimal System, 1942-2023

Dóra Pákozdi

1702 Salvor Tower, 41 Mastmaker Road, E14 9XS, London, UK
uczcpak@ucl.ac.uk

Dóra Pákozdi is a Hungarian-British autistic senior library assistant in a London public library while retraining in library studies at UCL, having previously completed another MA in comparative literature. She is an advocate for autistic liberation in both classification schemes and at workplaces.

Pákozdi, Dóra. 2024. "Sort of People: Considerations About the Ontogeny of Autism in the Dewey Decimal System, 1942-2023". *Knowledge Organization* 51, no. 8: 613-625. 78 references. DOI:10.5771/0943-7444-2024-8-613.

Abstract: *Sort of People: Considerations About the Ontogeny of Autism in the DDC, 1942-2023* traces the ontogeny of autism within the Dewey Decimal System, from the 14th to the 23rd editions. This period marks significant shifts in the psychiatric understanding and societal attitudes toward autism, reflecting the broader dynamics of neurodiversity recognition, the conceptualization of autism as a spectrum, and the influential role of autistic self-advocacy. The study draws on interdisciplinary sources and theoretical frameworks, including Ludwig's 'restricted malleability' and Hacking's 'interactive types' to critically analyse how changes in scientific, political, and social landscapes have influenced the organization of literature on autism in library classification systems. Methodologically, the paper employs a detailed historical review of DDC editions alongside an analysis of literature concerning the classification of various marginalized groups as well as medical, philosophical, and disability advocacy literature to map the shifts in autism's classification. This approach highlights how the language and structure of classification systems both reflect and shape societal attitudes towards autism. The analysis also considers the impact of autistic self-advocacy on challenging and reshaping these classifications, emphasizing the importance of language and representation in the struggle for autistic integration and visibility. The implications of this study extend beyond the specific case of autism classification. It contributes to ongoing debates on the politics of knowledge organization, the role of libraries and classification systems in either perpetuating or challenging societal norms, and the importance of incorporating marginalized perspectives in the creation and dissemination of knowledge. The author is diagnosed autistic.

Received: 29 June 2024; Accepted 31 July 2024.

Keywords: Ontogeny**;** autism; Dewey Decimal Classification; Knowledge organization systems.

"Not all of us can say, with any degree of certainty, that we have always been human, or that we are only that. Some of us are not even considered fully human now, let alone at previous moments of Western social, political and scientific history".

Rosi Braidotti, *The Posthuman*

1.0 Introduction

Language's dependence on context is made materially evident in subject headings. A long-standing classification scheme's aim is to collocate a type of resource based on interpretation of their domain (Tennis 2012). Hegemonic library classifications like the Dewey Decimal System have upheld and helped entrench marginalizing models of disability simply because they were created by certain cultural and scientific communities. Much of the ableism of the current age can be traced back to the Western eugenics movement of the 19th-century, when perceived physical, sensory and cognitive differences were first combined into the coherent biological other that is more of a product of normalcy than an excluded part of it (Snyder and Mitchell 2015). Shaped by scientific discovery and self-advocacy, autism's ontological status has been evolving, resulting in the classificatory confusion explored below.

1.1 Literature Review

There is a rich history of critique of the disciplinary structures and subject access standards shaping the classification of resources about marginalized groups both on the descriptive and on the structural level. Both the language used to label categories of literature about groups of people and the order in which the resources are arranged has inevitably affected how users and information professionals conceptualize individuals belonging to said group. In a school library organized by the Dewey Decimal System, an autistic teenager for example would have to search for books on mental health, self-help and practical life advice under 305.9085 'People with developmental disabilities' nested under 'People with disabilities and illnesses, gifted people' and 616.85882 'Autism Spectrum Disorder' collocated with books on an array of mental illnesses and – due to the small volume of school library holdings – medical resources. Abled, white, neurotypical, cisgender and straight students would look under 158.1 'Personal improvement and analysis'.

Radical cataloguer Berman's 1971 introduction to his foundational *Prejudices and Antipathies* summarily describes the users LC headings were created to satisfy:

> parochial, jingoistic Europeans and North Americans, white-hued, at least nominally Christian (and preferably Protestant) in faith, comfortably situated in the middle- and higher-income brackets, largely domiciled in suburbia, fundamentally loyal to the Established Order, and heavily imbued with the transcendent, incomparable glory of Western civilization. Further, it reflects a host of untenable-indeed, obsolete and arrogant-assumptions with respect to young people and women. And exudes something less than sympathy or even fairness toward organized labor and the sexually unorthodox or 'avant-garde' (Berman 1993,15).

Within classification systems, power is expressed through naming and ordering the phenomena of the world in relation to one another. Berman aimed to remedy this by correcting perceived biases within the LCSH, while other influential thinkers, like Hope A. Olson (Koford 2014) theorize that injustice is innate or inevitable in classification schemes. Members of groups falling outside Berman's description have since made tremendous efforts to represent themselves and their unique knowledge and ways of knowing: Duarte and Belarde-Lewis (2015) shows a vision of a world where the 600 indigenous tribes within US political borders each utilized their own knowledge organisation system based on their communities' unique epistemic customs, through analysing ten years of messages on two electronic cataloguing lists. Ho (2005) showed how most patrons' want to search foreign films by languages and that cataloguers support the addition of language and production form/genre headings to enhance their browsing experience, while Wagner (2022) explored body-oriented description as a new approach to cataloguing visual information in a gender inclusive way and Furner (2007) introduced critical race theory as a framework for evaluating classification schemes.

2.0 The Neurodiversity Framework and Library Classifications

The present work aims to trace the ontogeny of autism throughout editions 14 to 23 of the Dewey Decimal System. Much like intersexuality (Fox 2016) the subject of autism is particularly suitable to showcase breakdown and scatter in a classification scheme while the process of clamouring to accommodate the rapidly increasing volume and variety of publications that are in turn trying to make room for a previously invisible people and identity. This space – whether physically on a library shelf or theoretically in a controlled vocabulary – seems to be a scarce resource.

Tennis (2012) notes that the scatter caused by the inflexibility of an ableist classification scheme is the result of interpretations made by certain agents about literary or user warrant of the domain and not by changes of semantics altering collocation. The history of allistic medical experts and classification specialists having exclusively shaped the knowledge available about autistic people can be conceptualized through Wittgenstein's (2011) 'language games' where communities decide the agreed-upon terms through practical considerations. Mai (1999) argues that a knowledge organization system is a standardized version of these language games of a given organization and that a terminology fixed in such a way is highly political and can be both a great asset and a real a threat to freedom of expression. As autistic self-advocacy's visibility increases, the hope is that autistic people and our allies will be behind most of the knowledge disseminated about autism.

From a disability studies perspective, libraries can be thought of as locations of marginalization, and the social spaces where classification schemes were devised as links of a network of culture in which the divide between abled and disabled was already innate and while terminology had shifted from 'degenerate', 'defective, 'retarded' and 'deficient' to 'disordered', the structural underpinnings remain mostly unchanged. Despite librarians having addressed that schemes like the Dewey Decimal System (DDC), Library of Congress Classification (LCC), and the Library of Congress Subject Headings (LSCH) fail to accurately collocate resources about groups of people lacking political or economic power (Adler et al. 2017).

2.1 Classifying Autism: an Introduction

The classification of autism as a Dewey Decimal System subject follows the rapid advancement of psychiatry over the course of the 20th-century, trailing behind the similarly recent expansion of self-advocacy of neurodivergent humans. Drabinski (2013), critiquing the disciplinary structures shaping the cataloguing of queer library materials points out how librarians have failed to properly and respectfully catalogue works on marginalized groups, but instead of placing the strain on cataloguers, invites users to engage with catalogues as biased texts. Extant hegemonic classification schemes are resistant to change because of baked-in inertia (Bowker and Star 1999) while also causing significant scatter (Olson 2008). The uncertainty about autism's location in the totality of human knowledge has resulted in books about autism shelved in medicine, self-help, biographies, mental illness and parenting. Educating library users to critically evaluate and enhance these same systems that librarians work within and against has the potential of lessening the scatter and the classificatory marginalization of disabled people.

Nearly three decades after Leo Kanner's influential report on what he called Infantile Autism, research finally agreed to treat autism as an independent diagnosis of Schizophrenia, based on an overview of symptomatic differences like age of onset, family histories and responses to treatment (Meyer et al. 2011). Another three decades were needed for the Dewey Decimal System's 22nd edition to catch up with the advancements. Further complicating the diagnostic history and hence classification of literature about autism are contemporary concepts like the spectrum and neurodiversity, medical breakthroughs like the merging of autism and Asperger Syndrome diagnoses, and the rapidly increasing visibility of autistic self-advocacy. The debates and discussions ignited by self-advocacy are underpinned by the underlying framework of neurodiversity; concepts like the Double Empathy Problem (Milton 2012) or Monotropism (Murray et al. 2005) have had genuine impact in research, practice (Leadbitter et al. 2021) and the self-understanding of many neurodivergent and autistic people. The fundamental change is merely from a focus on normality to one on prevalence. Doyle (2020) proposes a new umbrella term for atypical neural and cognitive phenotypes: a neurominority is a population sharing a particular symptom cluster and encountering similar challenges in a majority neurotypical society. According to CDC data, while in 2002 one in every 150 US citizen was diagnosed autistic, by 2020 this number has expanded to one in every 36.

3.0 History of the concept of Autism

Literary scholar Murray's (2012) monograph on autism, according to its blurb, aims to "present a rounded portrayal of the ways in which autism is currently represented in the world" (42) yet takes a disappointingly reductive view of research into possible autistic individuals' lives, writing that to search for autism in the past is "a point about searching for definitions of it in the present" and calls for any claim about the pre-scientific history of autism "to be made with real expertise, and not left to excited or over-eager guesswork".[4] The reluctance to centre autistic people in our own history in favour of an ever-shifting, often disgraced cast of scientists is a major setback in the way of Autistic Liberation – the implication that only the professional and medical establishment is considered "real expertise" on identifying autism is rooted in the same confused, medicalised, pathology-focused concept of autism, enabled by library classification systems over the last century.

3.1 Autism, a Conceptual History: Beginnings

Prior to a scientific understanding of disability, neurodiversity and mental illness, many cultures conceptualized autistic people as changelings, elves or other almost-human creatures, tragic and distant, sometimes uniquely useful burdens on their communities (Wing 1997). An echo of this diversity of cognitive style and ability – seen by allistic people as seemingly contradictory strengths and weaknesses of autistic individuals – is present in the collocation of disability, illness and giftedness under the class number 305.908. There is enough of an abundance of autistic voices – with diverse support needs – in online autistic groups, the arts, media and our personal lives, to know that the autistic people of today still find ourselves in the depicted near-humanness of elves, robots and aliens, while an equally large group passionately dislikes the comparisons and the ableist implications they are riddled with. It doesn't require guesswork that previous generations followed similar lines of thought. A disabled child was often thought of as a monstrous fey creature corrupted or exchanged by harmful spirits (Wing 1997). This is an enduring concept exemplified by incidents as recent as Lord's (2006) award-winning autobiographical novel *Rules,* in which the mother of an autistic child muses "Sometimes I wish someone would invent a pill so David'd wake up one morning without autism, like someone waking from a long coma, and he'd say, "Jeez Catherine, Where have I been?" or the 2007 Ransom Notes Campaign – in which autism, depicted as an ominous, child-snatching villain, left threatening messages to the victims' distressed parents – that in many ways signalled the beginning of the Neurodiversity Movement (Kras 2010).

The momentum of enthusiasm for scientific classification originates from the early modern period and only ran out of steam during the early 20th-century, producing traditional classification schemes tending towards a concern primarily about an often false sense of objectivity and neu-

trality (Hjørland 1998). Since the advent of psychiatry and psychology different methods of classifications have been used to organise these scientific domains' knowledge about autism and atypical neurologies in general, each system reflecting the theoretical approach and the biases, views and personal horizons of their scientific authors. Knowledge about autism generated outside institutional or scientific settings was not taken into account by existing classification schemes, some of which now struggle to include it.

In her book, *Letters to My Weird Sisters: On Autism and Feminism*, autistic scholar Joanne Limburg takes a radical, 'dishuman' approach to bypass the medical establishment's monopoly on defining autism; she addresses personal letters to a neurominority: historic women who shared the same or similar lived experiences that she had had as an autistic cisgender woman. The 'dishuman' approach seeks to trouble established notions of what it means to be human based on Braidotti's concept of the posthuman (2013) from the stance of Disabled Liberaton.[3] Steve Silberman, author of *Neurotribes* was chastised by the *British Journal of Psychiatry* for identifying autism in historical figures and, according to the review's author, Lisa Conlan "It is hard to shake the feeling you are being toyed with in the name of a bigger political agenda". The agenda Conlan wants to avoid is Disabled Liberation.

Milton (2012) contests scientific methodologies' notion that the ability of neurotypical individuals to estimate the mental states, motivations and emotions of a peer is what constitutes empathy, when autistic people often see these estimates as inaccurate, hurtful or ignorant. The resulting communicational disconnect is rarer and often a brief interlude in the social and cognitive reality of neurotypicals, but is a traumatising everyday reality for autistic people. Hacking (2009) puts it: "There is a partial symmetry between the autistic and the non-autistic. Neither can see what the other is doing. The symmetry is only partial because we have an age-old language for describing what the non-autistic are feeling, thinking and so on, but are only creating one for the autistic" (1471). A 20-year-old non-speaking advocate, Noah Seback, after summarising how he considers most of his education 'warehousing' and 'babysitting' because of educators' refusal to imagine autistic students as agents, spelled out 'presuming competence' on his letterboard (Holmes 2024)[1].

3.2 Autism, a conceptual history: pathology

Although much less known than Leo Kanner's or Hans Asberger's reports, Soviet scientist Grunya Sukhareva was the first to pathologize what we now understand to be autism (Sher and Gibson 2023). Despite her description of autistic traits being remarkably close to those expressed in current diagnostic criteria her status as a Soviet woman under Stalin prevented her ideas from being disseminated outside of the Soviet Union (Manouilenko and Bejerot 2015).

Kanner's 1943 report examined two essential areas of autistic difference: social disconnectedness and a strong preference for sameness. He had also described autistic behaviours like echolalia and repetitive movements, and suggested that autism was innate and that autistic children often resemble their parents, a notion which under Bruno Bettelheim grew into the harmful idea that autism is caused by emotionally neglectful parenting, particularly perpetrated by mothers. Although the mid-century brought unprecedented upheaval to the field of psychiatry, Kanner's superficial association of autism and schizophrenia also hindered progress, with the 19th-century protocol of institutionalisation of non-speaking and intellectually disabled autistic individuals in inhuman, neglectful and abusive asylums (Donvan and Zucker, 2016) has been constant ever since; according to NHS England data 2045 autistic people and/or people with learning disabilities were inpatients in February of 2024 and 1075 of these patients had a length of stay over 2 years.

Autism was established as a separate diagnosis in 1980 after several competing definitions of autism had been proposed by Rutter (1978) and the The American National Society of Children (1978) and the comparisons to schizophrenia established autism as its own condition (Rutter 1972). Autism was entered into the DSM-III (APA 1980) under the new class of Pervasive Developmental Disorders and the conceptual realm of deficit, disorder and pathology. Kirk and Kutchins (1994) in their analysis of the lack of reliability of the DSM-III quote the manual stating how "It is particularly encouraging that the reliability for such categories as schizophrenia and major affective disorders is so high" (1980, 468) which is contradicted by the relationship between autism and schizophrenia being described as "controversial" by the new DSM article on Infantile Autism explaining: "Some believe that Infantile Autism is the earliest form of Schizophrenia, whereas others believe that they are two distinct conditions" (1980, 87).

The WHO's International Classification of Functioning, Disability and Health (ICF) was accepted by all the organisation's member states in 2001 as a comprehensive model and classification of disability that uses codes to capture the details of disabled individuals' functioning across various domains. The categories are flexible, but are often reduced to Core Sets – comprising only categories most pertinent for each condition – for clinical utility (Hayden-Evans et al. 2024). The International Classification of Diseases–Eleventh Revision (ICD-11) recommends clinicians use ICF categories to describe the impact of health conditions on individual functioning (Bölte 2018). The aim is to tailor services and resource-allocation to the individual needs of a disabled person, instead of providing potentially

mismatched, purely diagnosis-based care. The WHO Disability Assessment Schedule 2.0 is based on the ICF and was published in 2012, a year before the DSM-5 officially introduced the spectrum-model of autism.

Emphasizing variety over pathology, autistic self-advocates argued that disability is largely created by various barriers erected by ableist societies, as opposed to by some inherent deficit of autistic and neurodivergent humans. The spectrum-model of autism (Wing 1993) imagines autistic traits in diverse constellations, unique to each individual, resulting in a personal profile of support needs, cognitive styles, communication difficulties or preferences, and sensory differences – finally included in the DSM-V's diagnostic criteria. This more nuanced understanding of autism has led to the wave of diagnoses reported by the CDC above.

Despite the current fears about an 'autism epidemic' and 'overdiagnosis', studies have shown that early intervention improves developmental outcomes for autistic children, while caregivers are empowered to be able to support their families by accessing specialised support and community connections while reducing their own parental stress (Okoye et al. 2023). In order to reframe effectiveness in a neurodiversity-affirming way, early intervention has to centre well-being, autonomy, coping strategies and autistic-prioritised intervention targets (Leadbitter et al. 2021). In her article *What Can Physicians Learn from the Neurodiversity Movement?* Dr Christina Nicholaidis (2012), a mother of an autistic child urges her fellow medical professionals to "try to understand an individual's complex combinations of strengths and challenges, as well as the potential for wide variations in functioning". Instead of focusing on behaviours deemed impairments, autistic self-advocates seek to reframe the pathology model into one that addresses the difficulties of autistic humans informed by an Ethics of Care.

3.3 Towards neurodiversity

A useful narrative picture regarding the published usage of autism-related terms and synonyms can be derived through a Google Ngram Viewer search, showcasing instances of usages of terms within Google Books' digital corpus (Figure 1). The first occurrence of the term 'autism' coincide with Leo Kanner's mid-1940s works. His extremely narrow definition of autism can be seen in the minuscule dent 'autism' makes in the frequency of use of 'Schizophrenia', the previous only overtaking the latter a decade after the DSM-III established autism as a condition entirely separate of Schizophrenia and in the immediate aftermath of Lorna Wing's influential work popularizing the long-forgotten Hans Asperger (Wing 1981) and broadening of the concept of autism (Wing 1993). 'Infantile Autism' disappears at the same time, somewhere in the 1990s.

Despite the current fears about overdiagnosis, the evidence suggests that the largest spike in published interest in autism occurred between 2000 and 2010, not after the publication of the DSM-V. The extremely low rate of 'Neurodiversity' occurrences is a striking reminder of the discrepancies between the linguistic preferences of self-advocates versus the scientific world. The umbrella term 'Neurodiversity' was first coined by autistic sociologist Judy Singer in 1998 as a non-medical term that shifts the focus from deficit to the variety of human neurology and was quickly and enthusiastically embraced by the majority of the autistic community. It erases the need for binary medical or social models of autism. The above graph illustrates the scientific and publishing world's reluctance to follow suit.

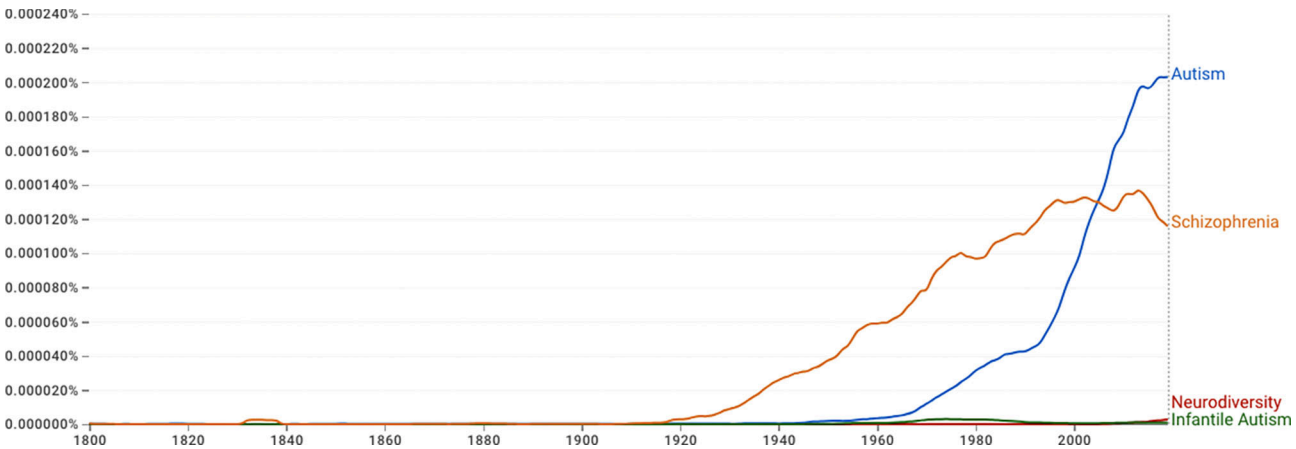


Figure 1. Google Ngram Viewer instances of competing terms to describe autism.

4.0 Theoretical considerations

A distinct characteristic of the Dewey Decimal System is its enumerative focus on disciplines. The same subject may appear under any discipline it has sufficient literary warrant in. This poses unique problems for the classification of disabilities and identities. To illustrate these problems, let's consider the example of the topic of autistic children's play. Lining and sorting toys in joyous solitude has been widely described as an early symptom of disordered development. Despite the hypothetical parents' goal to seek out activity books or collections of ideas to accommodate and support their child's style of play, they will be bombarded by lists pathologizing perfectly healthy autistic behaviours, creating confusion and distress. This is akin to classifying Anne of Green Gables under 'abnormalities' based on the rarity of natural red hair. Another example is that of unscientific memoirs by autistic people being routinely shelved and displayed in medical sections. Parents of autistic children looking for inspiring, empowering stories under 'First Facts' or 'First Experiences' style sections under class number 616 – common in UK public libraries – may find, nestled between infant health handbooks and specialist nutrition guides, Ma's (2017) picture books *In My World*, in which at the mention of autism, all colour drains out of the previously depicted images of adventure, family and fun and the reader suddenly finds the main character standing on a bleak, empty page inscribed "In your world I have autism".

4.1 Restricted malleability

Like Tennis' example of eugenics (2002) demonstrates, subjects can wonder in and out of disciplines as human interest and published scholarship change over time. Eugenics started out in the 1910s in the realms of Science, under the 500s near Genetics within Biology and have since moved into diverse classes of mostly Social Sciences. Over time, these shifts break down the collocative integrity of the scheme; books on the biological aspects of the human eugenics movements may sit next to plant reproduction. Autism in the Dewey Decimal System has a similarly shaky trajectory emerging from obscurity through the classes of 'Dementia' and 'Schizophrenia' to reach its own term by 2003 – but remaining mostly in the realm of medicine.

A useful concept for understanding this journey is 'restricted malleability' (Ludwig, 2023). It is similar to Ian Hacking's explanation of 'interactive kinds' (1999). Hacking's example is of another neurodiverse condition, ADHD. He argues that the interaction between the terms used to classify things or people and the people themselves are in a discourse, both affecting the other. People diagnosed autistic differ from non-speaking, disabled or eccentric humans of the past, in the ways they are diagnosed, treated and theorised about, but in being medicalised in such ways, autistic people had also had an effect – on the concept of disability and eccentricity.

Ludwig (2023) suggests a shift in focus from naturalness to materiality in debates about scientific classifications and argues that material kinds are affected by intervention. Neither the restriction nor the malleability is confined to either dimension of conceptual or non-conceptual, and one of Ludwig's examples is psychiatric kinds: these are restricted by the structure of their domain but also by their applicability – overly complex classifications are a hindrance to psychiatric practice. Ludwig adds that the malleability of kinds is a multi-dimensional gradient: categories like 'biological species', 'chemical elements' and 'bosons' are less malleable than classes of the likes of furniture, mental illness or genders and each of these can be shaped along underlying structures (see Figure 2).

Malleability	Material properties of 'autistic person' are shaped by changes in scientific advancements and classifications.	Classifications are shaped by changes in the material properties of 'autistic person'.
Example	The merging of previously separate diagnoses collocated 'people with Asperger Syndrome' and 'autistic people'.	Increased rates of diagnosed women and genderqueer people mean autism is not thought of as an exclusively male condition anymore.
Restriction	Material properties of 'autistic person' restrict and limit classificatory options.	Classifications are limited by demands towards linguistic representation.
Example	The lived experience of the sum-total of 'autistic people' should be the basis and entirety of what science considers 'autism'.	Example: autistic self-advocates demanding the removal of deficit- based language.

Figure 2. The restricted malleability model of the debate and progress of the classification of autism based on Ludwig's (2023) model.

4.2 Centring user needs

The exclusion of autistic people from the project of classifying their experiences and the resources about themselves is one of the underlying systems along which knowledge about autism has been organized. Since non-medical publications about autism and – knowingly -autistic researchers and cataloguers are a relatively new phenomenon, the literary warrant for 'autism' and its adjacent concepts have almost exclusively been based on medical texts. Fox (2016) demonstrates how in the case of a marginalized and poorly understood community, confusion in literature then provides the ontological basis for including terms and designing classifications, leading to even deeper marginalization. To alleviate the catastrophic levels of exclusion autistic people face in education, employment, healthcare and our social environment, this process needs to be reversed.

Rey (1995) argues that the selection of classificatory terms should depend on the explanatory work one wants concepts to perform his work must centre inclusion and an affirmative attitude towards the material realities of neurodivergent behaviours. While this pragmatism has its drawbacks – Fox' (2016) example of Intersexuality being classed under sexual orientation rather than the facet 'People by sex or gender' was a pragmatic decision based on the perceived users' needs as they are more likely to search for resources about Intersexuality near LGBTQ+ issues – controlled vocabularies could be harnessed for the work of Neurodivergent and Autistic Liberation.

5.0 14th-21st edition: Schizophrenia

The understanding of disability and attitudes towards its attendant non-normative behaviours varies greatly across time periods and cultures (Longmore 2003). Autistic traits like monotropic attention and advanced pattern recognition were sometimes highly desirable skills that helped some privileged disabled individuals to lead fulfilling and happy lives. Henry Cavendish, described as "the coldest and most indifferent of mortals" – could calculate the density of the Earth from his 18th-century home while living in near-perfect isolation and struggling to make eye-contact (Silberman 2016). Albeit community norms about acceptable behaviours have always existed – Cavendish was thought beyond eccentric when, upon encountering his maid one day, had a separate staircase built in order to avoid all human interaction (Silberman 2016) – scientific interest in pathologizing behaviour began only when psychology first distinguished itself from philosophy.

5.1 14th Edition

The 14th edition, of the Dewey Decimal System, published in 1942, a year before Kanner's (1943= landmark report, carries a fossil from those times by classifying Schizophrenia both at 132.1982, and at 616.8982 under 'Dementia precox and schizophrenic reaction types'. The classmark 132.1 was for 'Insanity' and 'Mental alienation' while 616.8 stood for 'Diseases of nervous sistem' and 'Psychiatry'. During this time, in 1943, Donald Triplett was diagnosed as "Case 1" of autism by Leo Kanner. Triplett was US American, white, male, had savant abilities and was born into an affluent bank-owning family (Pallardy 2024). Temple Grandin – also white US American and affluent with savant abilities – was diagnosed in 1950, opening up the possibility for women to be considered autistic (Richter, 2014).

5.2 15th Edition

The next, 1951 edition shows the mid-century acceleration of psychiatric progress by simplifying the class and listing the diagnosis and treatment of conditions like Schizophrenia, Paranoia and Manic-Depressive Psychoses under one classmark at 616.89. 616.8 is also referenced in the notes of the 132 class titled 'Abnormal psychology'. These read "*Includes irrational, abnormal, or deranged mental processes; their causes; mental symptoms of disease*" and "*For Medical treatment of these disorders, see Nervous system and neurology, 616.8; Psychoses and psychiatry, 616.89*". It appears to demarcate 'Abnormal psychology' under 132 and 'Psychology' under 150. Clearly, those behind this decision had a precise work they wanted these concepts to perform; however intentional, this change has the clear impact of marking the neurology and behaviours of some humans as flawed and unnatural. Since then, 'deranged mental processes' like autistic humas' need for sameness for example were proven to be linked with the anxiety and dread most autistic people feel when encountering unexpected change (Uljarević et al. 2017). Longmore (1985) demonstrates how disabled people are often viewed as mere objects of medical attention even far removed from medical settings. A consequence of this is the picture of the adult disabled patient as impaired, docile, childlike and eternally grateful to be at the mercy of medical professionals. Any autistic adult searching the internet or libraries for advice is intimately familiar with being bombarded with resources addressing exclusively allistic parents of autistic children.

The DSM's first edition was published in 1952, simplifying the confused and varied documents used by psychiatric professionals for diagnostic practice. Since its inception in the mid-twentieth century, the Diagnostic and Statistical Manual of the American Psychiatric Association – currently in its DSM-5-TR edition – has exerted tremendous

medical and cultural influence. By providing comparable samples, the DSM helped accelerate related research. The DSM is a principal guide to psychiatry and is used by policymakers, insurance companies, researchers and some psychiatric professionals. When research had cleaned up the lack of validity about autism as an independent diagnosis, in 1980 it was entered into the DSM-III with a heavy focus on its 'infantile' nature (Volkmar and Reichow 2013).

5.3 16th Edition

The 16th edition from 1958 seems the most confused. It introduces the term 'mental illness' and superimposes it on class 132.1 located next to 132.2 'mental deficiency', the class containing topics like drug addiction and queer sexualities. Between them it also establishes a 'functional psychoses' group at 132.19 containing many types of Dementia, Schizophrenia and Manic-depressive Psychoses. This is also the first edition to mark separate classes in the 300s for services related to mental illnesses and disabilities. The medical class number for Schizophrenia remained unchanged here and in the next three edition as well with the difference of classes between 616.892-616.898 labelled 'specific psychoses'.

5.4 17th Edition

In the 17th edition of 1965, the domain of increasingly professionalizing psychology resulted in tighter collocation and the domain was organized under 150, with 130 being left behind to represent 'Pseudo-and parapsychology'. The 157s were dedicated to 'Abnormal and clinical psychologies' having previously shared the classmark 132 with the newly unscientific knowledge now under the 130s. 616.85-86 was reserved for 'Psychoneuroses' like 'Hysteria' and 'War Neurosis', Anxiety, Phobias, OCD and Epilepsy. The psychiatric domain was charted under 616.89 and seemed to contain only Schizophrenia under 616.8982 – then understood as a type of Dementia' and Dementia under 616.8983. Immediately following, 616.9 was dedicated for 'Communicable diseases'. It is immeasurable how much damage was caused by collocations like this encouraging the public to associate various neurological differences and impairments with transferable diseases.

The term 'autism epidemic' remains in use by scientifically illiterate journalists who often deliberately or ignorantly misrepresent data about autism to garner a response of shock and concern. Hill's (2024) Guardian article *What's behind the UK's increase in autism diagnoses?* writes that one of the reasons behind the increase is that in the 1980s only a quarter of intellectually disabled people were also diagnosed autistic. Hill (2024) writes "Now the NHS acknowledges that it could be as high as three-quarters." The citation clickable on the words "NHS acknowledges" links to a 2021 news article about the NHS's Long Term Plan, more precisely about how three-quarters of intellectually disabled people aged 14 and older have received an annual health check two year ahead of targets laid out in the Long Term Plan. To round the paragraph off, Hill quotes the multiply-discredited Baron-Cohen saying "That's an incredibly steep rise". The source of this quote remains unknown. Amelia Hill has written several widely-read articles about autism.

5.5 18th Edition

The accelerating progress and social capital of psychiatry and medicine can be seen mirrored in how the 18th edition further intertwines the two broad classes by instructing the classifier to "*Add to 157.2 the numbers following 616.89 in 616.895-616.898, e.g., manic-depressive psychoses 157.25*". Non-medical books about Schizophrenia and autism then would sit under 157.282.

5.6 19th Edition

The 19th edition in 1979 instructs classifiers to place resources about both organic and functional psychoses under the medical class numbers in 616.892-616.898, further shrinking the 157 'Abnormal and clinical psychologies' group. Schizophrenia remained under 616.8982 which remains a common class number for books about autism.

5.7 20th Edition

The 20th edition, published in 1989 is the first one to mention the term 'autism', albeit not with its own number. This came after 9 years of the DSM-III's establishment of an independent diagnosis and few years into Wing's (1993) research that would establish the spectrum-model of autism. In the medical schedules it is wedged into the same class as Schizophrenia, 616.8982 and in Education it is placed at 371.94 where autistic students serve as an example of 'emotionally disturbed students'. The notes even make sure to warn information professionals not to class all resources about 'delinquent and problem students' under this number. It is also worth noting how 'autistic students' is nestled between various degrees of 'retarded' and 'gifted' students. This placement hints at what autistic self- advocates call a 'spiky' or 'uneven profile'; a result of atypical neurological connectivity. In his literature review of EEG and MEG studies of autistic brains, O'Reilly et al. (2017) identified a clear trend of long-range functional underconnectivity. This explains why autistic individuals might struggle with some seemingly basic tasks, while excelling at areas requiring deep attention and pattern-recognition. Atypical strategies for

the allocation of attention are employed by most Neurodivergent people and the framework of Monotropism (Murray et al. 2005) can help demystify autistic cognitive styles by deconstructing myths and misunderstandings about autistic life.

5.8 21st Edition

By the time of the 21st edition, published in 1996, Asperger Syndrome was formalized into a short-lived diagnosis and a Dewey Decimal System class under 616.858832. This has only recently been depreciated and then discontinued during the summer of 2023, ten years after the current DSM-V had merged the two into 'Autism Spectrum Disorder'. This delay probably contributed to the immense confusion about merging of the two diagnoses.

6.0 22nd-23rd edition: Autism

The 2000s were a turbulent time for autism-related research, with Andrew Wakefield's anti-vaccination fraud stretching from the 1998 publication of the since retracted paper on a debunked causal link between MMR vaccines and autism to the Lancet's 2010 retraction of it. In 2003 Simon Baron-Cohen's published *The Essential Difference: Men, Women and the Extreme Male Brain* in which he argues – among many disturbing and discredited findings – that autism is caused by 'extreme male brains'. It is not (van Ejk and Zietsch 2021). Around the year 2000, as seen on the above Google NGram diagram, the frequency of published mentions of autism dramatically surpasses that of Schizophrenia for the first time, indicating a dramatic increase in published material.

6.1 22nd Edition

2003's 22nd edition of the Dewey Decimal System was the first to include a separate class for the medical aspect of autism. This edition discontinued to use of 616.898 for autism and reshuffled the classification of Schizophrenia. The former was relocated to 616.85882 while the latter was moved mostly to 616.8581 where it is more tightly collocated with various personality disorders – albeit still within a grab-bag of 'Personality disorders, sexual disorders, impulse-control disorders, factitious disorders, developmental disorders, learning disorders; violent behaviour; intellectual disabilities' under 616.858. Medical, therapeutic and etiological resources remained under 616.898. Implementation of all these subtle changed across various institutions must have taken longer.

6.2 23rd Edition

Currently autism sits under 616.85882, immediately under '*Intellectual disabilities; developmental and learning disorders*', a category that neighbours a slew of extremely varied conditions and neurological variations from ADHD through APD and even homicidal behaviour. It also neighbours an intensely subdivided class of 'other congenital abnormalities' like Hydrocephalus and in public libraries with smaller collections, patrons might find themselves browsing for books about their own autism next to books about cancer or animal husbandry.

A welcome improvement is the acknowledgement of the marginalization of autistic people, and more especially autistic adults. Although class numbers for non-medical resources about autism are a bit more difficult to find. With the addition of facets, 362.196-362.198:001-009 'Services to patients with specific conditions' can be used for resources about social services and 305.9085 for works on autistic people ourselves. Removed from the realm of medicine, it is a welcome change to see neither of these containing terms like 'disorder'. The regrettable (over)use of terms like these is comparable to Fox's (2016) study showcasing how Intersexuality used to be classed under 'monstrosities' and 'sexual diseases'. A further, non-invasive improvement could be to use 'condition' in place of 'disorder'[2] and leave the notion of normality behind.

7.0 Conclusion

Broadly speaking, the concept of autism has emerged from the total obscurity of mythic monstrosity and benign, if shunned, eccentricity, into the confused pair conditions of schizophrenia/ autism to split into two variations around perceived functioning levels and is now slowly becoming a group of people with personal profiles of a constellation of many autistic traits and abilities. Evident in this process are the competing concepts of variation versus pathology, the latter dragging with itself a long history of epistemic objectification and outright threat of extinction, but also the very real expression of barriers preventing autistic people from living fulfilling and happy lives.

Martin (2021, 294) writes that "treating people and groups with respect necessitates calling them by the names they use for themselves". Autistic activist Bonello's (2022) research included interviews with 11,211 responders, most of whom were autistic that showed that despite terms like 'disorder' following autistic people everywhere they go – including libraries – most of us don't want to be called a disorder (Figure 3). Goldberg's (2023) analysis illuminates the dialectic of malleable social norms and neurodiversity through the example of left-handedness: it is caused by atypical brain lateralisation and left-handed individuals have his-

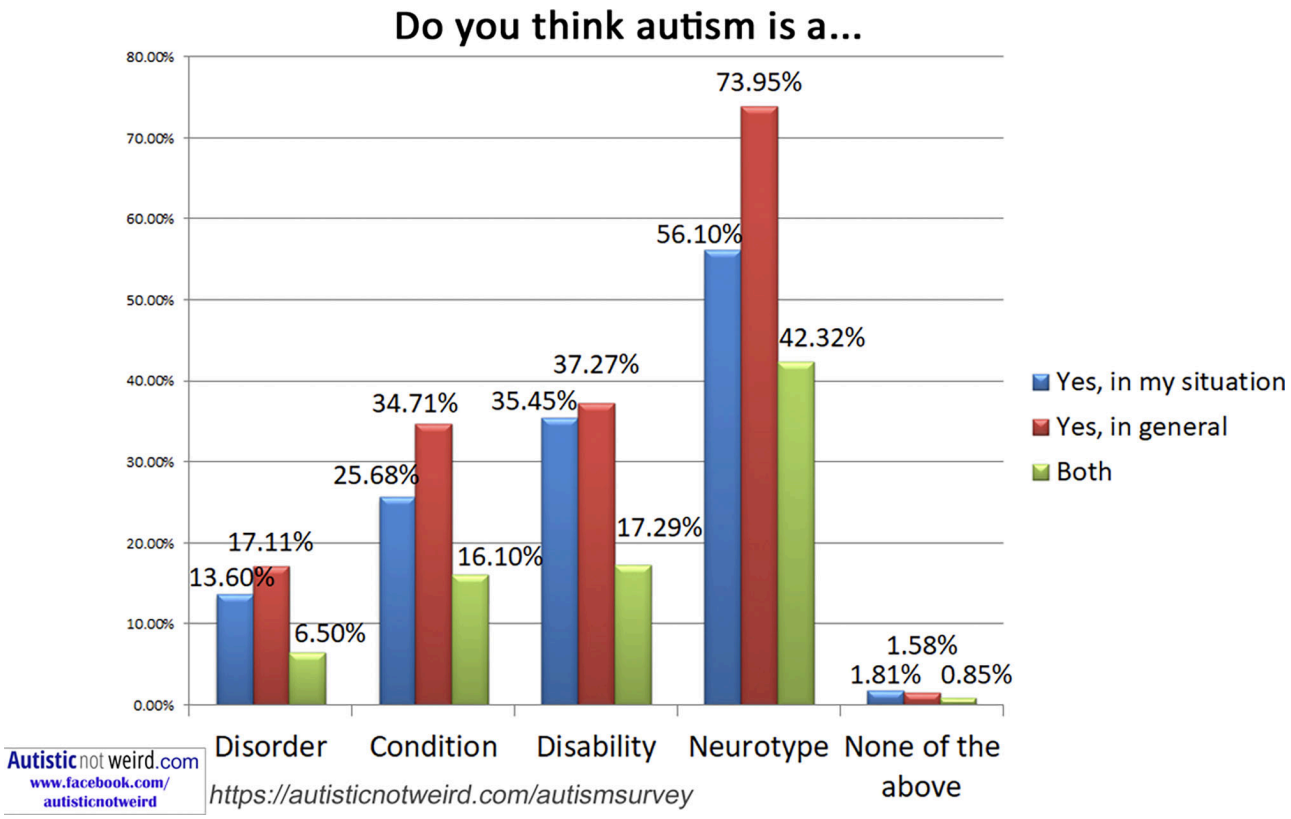


Figure 3. Chris Bonello's research showcasing preferences for autistic self-identification.

torically been discriminated against through lack of accommodations and through association with ritual impurity and undesirable qualities. Using accurate classificatory terms – placed correctly in their context – for autism would be a part of a similar normalisation process; instead of being called 'unclean', the term 'left-handed' bypasses the stigma and helps facilitate the necessary co-operation of affected individuals and providers of pertinent services. As an autistic public librarian, it is beyond jarring having to walk confused parents past bays of books listing every medical diagnosis from the bubonic plague to brain damage. Not shelving all – or the majority of – books about autism on medical shelves would help de-pathologise atypical neurology and benefit all humankind by encouraging library-users to stop othering autistic people.

On the subject of outdated terms in controlled vocabularies about queer people, Drabinski (2013) argues that trying to correct all of the terms is a secondary concern, if not outright dangerous as identities are deeply embedded in their cultural and chronological context and ripping them out of the reach of library users is destructive. The main difference between queer and autistic people in this matter however is that while the social aspects of autism has indeed changed over time, unlike queerness, autism is neurological: it results from atypical neurological connectivity. With the exception of the medical gatekeeping of gender- affirming care, queerness is no longer medically pathologized in most of the English-speaking world, but it doesn't mean that specialized services and spaces don't exist for queer people. A gay teenager doesn't need a brain scan or a multi-occasion diagnostic process to access identity-affirming care. Drabinski's (2013) approach – projected to autistic people's position in classification schemes – steers towards the marginalization of materiality, while providing ample space for the subjects themselves to assert their presence in the catalogue. She is however less interested in finding correct terms and instead asks "*Why don't I see myself in the subject vocabulary, and what does this tell me about the other ways I feel invisible*?". Anderson (2021) answers the above by conducting research in which autistic participants select their preferred way to communicate.

Another way to answer 'What is autistic invisibility like?' is based on this review of the ontogeny of autism in the ten most recent editions of the Dewey Decimal System: as agents and even as a unique people, we are at the cusp of leaving complete invisibility. The confusion of the knowledge domain is reflected throughout its literature, which then either crystallizes into confused and ever-changing classification schemes (Fox 2016) or becomes increasingly un-scientific. While the Dewey Decimal System's solutions

are less than ideal, its editors chose to follow scientific advancement, a jarring example of the opposite approach is NLM's treatment of autistic people: at WM 203.5. there is still no individual classmark for autism. This is the complete annihilation of autistic people's epistemic agency. Considering that many in the medical establishment still seek to 'cure' autism, this blank space is especially atrocious.

The author is diagnosed autistic.

Endnotes

1. See Noah's website at https://www.thisismenoah.com
2. A common practice in talking about Neurodivergent students in British secondary schools.

References

Adler, Melissa, Jeffrey T. Huber and A. Tyler Nix. 2017. "Stigmatizing Disability: Library Classifications and the Marking and Marginalization of Books about People with Disabilities". *The Library Quarterly 87*, no. 2: 117–35. https://doi.org/10.1086/690734

Alissa. 2021. "Classifying Autism spectrum disorders in Dewey Decimal System, LCC and NLM". *Cataloguing the Universe* (blog). October 26, 2021. https://lissertations.net/post/1758

APA (American Psychiatric Association). 1980. *Diagnostic and statistical manual of mental disorders.* 3rd ed. Washington, DC: APA.

APA (American Psychiatric Association). 1987. *Diagnostic and statistical manual of mental disorders* . 3rd ed., rev. Washington, DC: APA.

APA (American Psychiatric Association). 1994. *Diagnostic and statistical manual of mental disorders.* 4th ed. Washington, DC: APA.

APA (American Psychiatric Association). 2000. *Diagnostic and statistical manual of mental disorders*. 4th ed, rev. Washington, DC: APA.

APA (American Psychiatric Association). 2013. *Diagnostic and statistical manual of mental disorders.* 5th ed. Washington, DC: APA.

Anderson, Amelia M. 2021. "Exploring the Workforce Experiences of Autistic Librarians Through Accessible and Participatory Approaches". *Library & Information Science Research* 43, no. 2. https://doi.org/10.1016/j.lisr.2021.101088

Berman, Sanford. 1971. *Prejudices and Antipathies*: *A Tract on the LC Subject Heads Concerning People*. Metuchen, NJ: Scarecrow Press.

Bölte, Sven *et al.* 2018. "The Gestalt of Functioning in Autism Spectrum Disorder: Results of the International Conference to Develop Final Consensus International Classification of Functioning, Disability and Health Core Sets". *Autism* 23, no. 2: 449–67. https://doi.org/10.1177/1362361318755522

Bonello, Chris. 2022. "Results and Analysis of the Autistic Not Weird 2022 Autism Survey". *Autistic Not Weird*. Accessed 9 August 2024. https://autisticnotweird.com/autismsurvey

Bowker, Geoffrey C., and Susan Leigh Star. 1999. *Sorting Things Out: Classification and Its Consequences*. Cambridge, MA: The MIT Press. https://doi.org/10.7551/mitpress/6352.001.0001

Braidotti, Rosi. 2013. *The posthuman*. Cambridge: Polity Press.

CDC (Centers for Disease Control and Prevention). 2023. *Data and statistics on Autism spectrum disorder*. Accessed 9 August 2024. https://www.cdc.gov/autism/data-research/index.html

Dewey, Melvil. 1942. *Decimal classification and relativ index*. 14th ed. Lake Placid Club, NY: Forest Press Inc.

Dewey, Melvil. 1951. *Decimal classification and relative index.* 15th ed. Lake Placid Club, NY: Forest Press Inc.

Dewey, Melvil. 1958. *Decimal classification & relative index.* 16th ed., vol. 1. Lake Placid Club, NY: Forest Press Inc.

Dewey, M. 1965. *Decimal Classification and Relative Index.* 17th ed., vols. 1–2) Lake Placid Club, NY: Forest Press Inc.

Dewey, M. 1971. *Decimal Classification and Relative Index.* 18th ed., vol. 2. Lake Placid Club, NY: Forest Press Inc.

Dewey, M. 1979. *Decimal Classification and Relative Index*. 19th ed., vols. 1–2. Lake Placid Club, NY: Forest Press Inc.

Dewey, M. 1989. *Decimal Classification and Relative Index*. 20th ed. Lake Placid Club, NY: Forest Press Inc.

Donvan, John and Caren Zucker. 2017. *In a Different Key: The Story of Autism*. London: Penguin Books.

Doyle, N. 2020. "Neurodiversity at Work: A Biopsychosocial Model and the Impact on Working Adults." *British Medical Bulletin* 135, no.1: 108–25. https://doi.org/10.1093/bmb/ldaa021

Drabinski, Emily. 2013. "Queering the Catalog: Queer Theory and the Politics of Correction". *The Library Quarterly* 83, no. 2: 94–111. https://doi.org/10.1086/669547

Duarte, Marisa Elena and Miranda Belarde-Lewis. 2015. "Imagining: Creating Spaces for Indigenous Ontologies". *Cataloging & Classification Quarterly* 53, no. 5-6: 677–702. https://doi.org/10.1080/01639374.2015.1018396

Fox, Melodie J. 2016. "Subjects in Doubt: The Ontogeny of Intersex in the Dewey Decimal Classification". *Knowledge Organization* 43, no. 8: 581–93. https://doi.org/10.5771/0943-7444-2016-8-581

Frances, Allen. 2014. *Saving Normal: An Insider's Revolt Against Out-Of-Control Psychiatric Diagnosis, DSM-5,*

Big pharma, and the Medicalization of Ordinary Life. Boston: Mariner Books.

Furner, Jonathan. 2007. "Dewey Deracialized: A Critical Race-Theoretic Perspective". *Knowledge Organization* 34, no. 3: 144–68. https://doi.org/10.5771/0943-7444-2007-3-144

Ganin, Netanel. 2021 September 17. "On Adding Autistic People to HV". *I Never Metadata I Didn't Like* (blog). September 17, 2021. https://inevermetadataididntlike.wordpress.com/

Goldberg, Hagar. 2023. "Unraveling Neurodiversity: Insights from Neuroscientific Perspectives". *Encyclopedia* 3 no. 3: 972–80. https://doi.org/10.3390/encyclopedia3030070

Hacking, Ian. 1999. *The Social Construction of What?*. Cambridge, MA: Harvard University Press.

Hacking, Ian. 2009. "Autistic Autobiography". *Philosophical Transactions of the Royal Society B: Biological Sciences* 364: 1467–73. https://doi.org/10.1098/rstb.2008.0329

Hayden-Evans, Maya, Kiah Evans, Benjamin Milbourn, Emily D'Arcy, Angela Chamberlain, Bahareh Afsharnejad, Andrew Whitehouse, Sven Bölte and Sonya Girdler. 2024. "Validating the International Classification of Functioning, Disability and Health Core Sets for Autism in a Sample of Australian School-Aged Children on the Spectrum." *Journal of Autism and Developmental Disorders*. https://doi.org/10.1007/s10803-024-06295-5

Hill, Amelia. 2024. "What's behind the UK's increase in autism diagnoses?". *The Guardian*, March 4, 2024. https://www.theguardian.com/society/2024/mar/04/uk-increase-autism-diagnoses-neurodiversity#:~:text=That%20has%20gradually%20increased%20and,who%20can%20make%20the%20diagnosis.

Hjørland, Birger. 1998. "The Classification of Psychology: A Case Study in the Classification of a Knowledge Field". *Knowledge Organization* 25, no. 4: 162-201.

Ho, Jeanette. 2005. "Applying Form/Genre Headings to Foreign Films: A Summary of AUTOCAT and OLAC-LIST Discussions". *Cataloging & Classification Quarterly* 40, no. 2: 73–88. https://doi.org/10.1300/j104v40n02_06

Holmes, Stephanie C. (2024). "Noah's Voice: Non-Speaking Man Advocates for Big Changes". *Autism Parenting Magazine*, 26 Jan. 2024 (website). https://www.autismparentingmagazine.com/noah-voice-man-advocates-changes/#:~:text=Nadine%20and%20David%20Seback%20are,nonverbal%20individuals%20on%20the%20spectrum

Kanner, Leo. 1943. "Autistic Disturbances of Affective Contact". *Nervous Child* 2, no. 3: 217–50.

Kirk, Stuart A. and Herb Kutchins. 1994. "The Myth of the Reliability of DSM". *The Journal of Mind and Behavior* 15, no. 1-2: 71-86. https://www.jstor.org/stable/43853633

Koford, Amelia. 2014. "How Disability Studies Scholars Interact with Subject Headings". *Cataloging & Classification Quarterly* 52, no. 4: 388–411. https://doi.org/10.1080/01639374.2014.891288

Kras, Joseph F. 2010. "The 'Ransom Notes' Affair: When the Neurodiversity Movement Came of Age". *Disability Studies Quarterly* 30, no. 1. https://dsq-sds.org/index.php/dsq/article/view/1065/1254

Leadbitter, Kathy, Karen Leneh Buckle, Ceri Ellis and Martijn Dekker. 2021. "Autistic Self-Advocacy and the Neurodiversity Movement: Implications for Autism Early Intervention Research and Practice". *Frontiers in Psychology* 12. https://doi.org/10.3389/fpsyg.2021.635690

Longmore, Paul K. 1985. "A Note on Language and the Social Identity of Disabled People". *American Behavioral Scientist* 28, no. 3: 419–23. https://doi.org/10.1177/000276485028003009

Limburg, Joanne. 2021. *Letters to My Weird Sisters: On Autism and Feminism*. London: Atlantic Books.

Longmore, Paul K. 2003. *Why I Burned My Book and Other Essays on Disability*. Philadelphia: Temple University Press.

Lord, Cynthia. 2006. *Rules*. London: Scholastic Press.

Ludwig, David. 2023. "From Naturalness to Materiality: Reimagining Philosophy of Scientific Classification". *European Journal for Philosophy of Science* 13: 1-23. https://doi.org/10.1007/s13194-023-00509-w

Ma, Jillian. 2017. *In my world*. Illustrated by Mimi Chao. Arlington: Future Horizons Inc.

Mai, Jens-Erik. 1999. "A Postmodern Theory of Knowledge Organization". In *Knowledge, Creation, Organization and Use: Proceedings of the ASIS Annual Meeting Oct. 31-Nov. 4 1999, Washington, DC, USA,* edited by Marjorie M. K. Hlava and Larry Woods. Medford, NJ.: Information Today, 547-56.

Manouilenko, Irina and Susanne Bejerot. 2015. "Sukhareva – Prior to Asperger and Kanner". *Nordic Journal of Psychiatry* 69, no. 6: 479-82. https://doi.org/10.3109/08039488.2015.1005022

Martin, Jennifer M. 2021. "Records, Responsibility, and Power: An Overview of Cataloging Ethics". *Cataloging & Classification Quarterly* 59, no. 2-3: 281–304. https://doi.org/10.1080/01639374.2020.1871458

Meyer, Urs, Joram Feldon and Olaf Dammann. 2011. "Schizophrenia and Autism: Both Shared and Disorder-Specific Pathogenesis Via Perinatal Inflammation?". *Pediatric Research* 69: 26-33. https://doi.org/10.1203/pdr.0b013e318212c196

Milton, Damian E. M. 2012. "On the Ontological Status of Autism: the 'Double Empathy Problem'". *Disability & Society* 27, no. 6: 883–7. https://doi.org/10.1080/09687599.2012.710008

Murray, Dinah, Mike Lesser and Wendy Lawson. 2005. "Attention, Monotropism and The Diagnostic Criteria for Autism." *Autism* 9, no. 2: 139–56. https://doi.org/10.1177/1362361305051398

Murray, Stuart. 2012. *Autism*. London: Routledge.

National Society for Autistic Children. 1978. "Definition of the Syndrome of Autism". *Journal of Autism and Childhood Schizophrenia* 8: 162–7. https://doi.org/10.1007/bf01537864

NHS England. 2021. "Three in four people with a learning disability receive NHS annual health check". *NHS England, 10 Jun. 2021* (website). https://www.england.nhs.uk/2021/06/three-in-four-people-with-a-learning-disability-receive-nhs-annual-health-check/

Nicolaidis, Christina. 2012. "What Can Physicians Learn from the Neurodiversity Movement?" *Virtual Mentor* 14, no.6: 503–10. https://doi.org/10.1001/virtualmentor.2012.14.6.oped1-1206

Okoye, Chiugo, Chidi M. Obialo-Ibeawuchi, Omobolanle A. Obajeun, Sarosh Sarwar, Christine Tawfik, Madeeha Subhan Waleed, Asad Ullah Wasim, Iman Mohamoud, Adebola Y. Afolayan and Rheiner N. Mbaezue. 2023. "Early Diagnosis of Autism Spectrum Disorder: A Review and Analysis of the Risks and Benefits". *Cureus* 15, no. 8. https://doi.org/10.7759/cureus.43226

Olson, Hope A. 2008. "How We Construct Subjects: A Feminist Analysis". *Library Trends* 56. no. 2: 509–41. https://doi.org/10.1353/lib.2008.0007

O'Reilly, Christian, John D. Lewis and Mayada Elsabbagh. 2017. "Is Functional Brain Connectivity Atypical in Autism? A Systematic Review of EEG and MEG Studies". *Plos One* 12, no. 5. https://doi.org/10.1371/journal.pone.0175870

Pallardy, Richard, 2024. "Donald Triplett: American Autism Patient". *Encyclopædia Britannica* (website). https://www.britannica.com/science/autism-spectrum-disorder

Rey, Georges. 1995. "Concepts". In *A Companion to the Philosophy of Mind* edited by Samuel Guttenplan, 185–93. Oxford: Blackwell.

Richter, Ruthnann (2014). "5 Questions: Temple Grandin discusses autism, animal communication". *Stanford Medicine News Center*, 13 Nov. 2014 (website). https://med.stanford.edu/news/all-news/2014/11/5-questions--temple-grandin-discusses-autism--animal-communicati.html

Rutter, Michael. 1972. "Childhood Schizophrenia Reconsidered". *Journal of Autism and Childhood Schizophrenia* 2: 315–37. https://doi.org/10.1007/BF01537622

Rutter, Michael. 1978. "Diagnosis and Definition of Childhood Autism". *Journal of Autism and Childhood Schizophrenia* 8: 139–61. https://doi.org/10.1007/bf01537863

Sher, David Ariel and Jenny L. Gibson. 2021. "Pioneering, prodigious and perspicacious: Grunya Efimovna Sukhareva's life and contribution to conceptualising Autism and Schizophrenia". *European Child & Adolescent Psychiatry* 32, no. 3: 475-90. https://doi.org/10.1007/s00787-021-01875-7.

Silberman, Steve. 2016. *NeuroTribes*. London: Atlantic.

Singer, Judy. 2017. *Neurodiversity: The Birth of an Idea.* Self-published.

Snyder, Sharon L. and David T. Mitchell. 2015. *Cultural locations of disability*. Chicago: The University Of Chicago Press.

Tennis, Joseph T. 2002. "Subject Ontogeny: Subject Access through Time and the Dimensionality of Classification." In *Challenges in Knowledge Representation and Organization for the 21st Century: Integration of Knowledge across Boundaries: Proceedings of the Seventh International ISKO Conference 10-13 July 2002 Granada, Spain*, edited by Maria J. López-Huertas, and Francisco J. Muñoz-Férnandez. Advances in Knowledge Organization 8. Würzberg: Ergon Verlag, 54-9.

Tennis, Joseph T. 2012. "The Strange Case of Eugenics: A Subject's Ontogeny in a Long-lived Classification Scheme and the Question of Collocative Integrity". *Journal of the American Society for Information Science* 63, no. 7: 1350-9. https://doi.org/10.1002/asi.22686

Uljarević, Mirko, Amanda L. Richdale, David W. Evans, Ru Ying Cai and Susan R. Leekam. 2017. "Interrelationship Between Insistence on Sameness, Effortful Control and Anxiety in Adolescents and Young Adults with Autism Spectrum Disorder (ASD)". *Molecular Autism* 8: 1-6. https://doi.org/10.1186/s13229-017-0158-4

Van Eijk, Liza and Brendan P. Zietsch. 2021. "Testing The Extreme Male Brain Hypothesis: Is Autism Spectrum Disorder Associated With a More Male-typical Brain?". *Autism Research* 14, no. 8: 1597-608. https://doi.org/10.1002/aur.2537

Volkmar, Fred R. and Brian Reichow. 2013. "Autism in DSM-5: Progress and Challenges". *Molecular Autism* 4: 1-6. https://doi.org/10.1186/2040-2392-4-13

Wagner, T. 2022. "Body-Oriented Cataloging as a Method of Inclusive Gender Representation." *Cataloging & Classification Quarterly* 60, nos.6–7: 620–45. https://doi.org/10.1080/01639374.2022.2089795

Wing, Lorna. 1981. "Asperger's Syndrome: A Clinical Account". *Psychological Medicine* 11, no. 1: 115–29. https://doi.org/10.1017/s0033291700053332

Wing, Lorna. 1993. "The Definition and Prevalence of Autism: A Review". *European Child & Adolescent Psychiatry* 2: 61–74. https://doi.org/10.1007/bf02098832

Wing, Lorna. 1997. "The History of Ideas on Autism: Legends, Myths and Reality". *Autism* 1, no 1: 13–23. https://doi.org/10.1177/1362361397011004

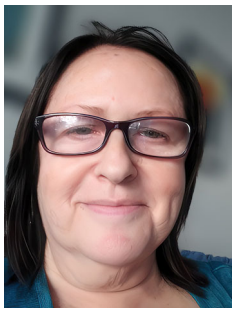
Wittgenstein, Ludwig. 2011. *Philosophical investigations*. Oxford: Wiley-Blackwell.

Native Peoples and Knowledge Organization: Perspective from the Indigenous Subject Representation to Promote Latin American Approaches

Mario Barité* and Mirtha Rauch**

* ** University of the Republic, Uruguay. Address: San Salvador 1944, CP 11200, Montevideo, Uruguay
*mario.barite@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.



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, and Mirtha Rauch. 2024. "Native Peoples and Knowledge Organization: Perspective from the Indigenous Subject Representation to Promote Latin American Approaches". *Knowledge Organization* 51, no. 8: 626-641. 81 references. DOI:10.5771/0943-7444-2024-8-626.

Abstract: The situation of many indigenous cultures in Australia, North, Central, and South America can be described as one of marginalization or minorization. Subject representation of Indigenous knowledge constitutes one of the contemporary crossroads since, through it, the predominant mentalities of classificationists, classifiers, and indexers are revealed, and this can consolidate hegemonic visions or propose appropriate alternatives to the cultural particularities of Indigenous peoples. From a critical perspective, this work aims to contribute to the systematization of the growing literature on indigenous warrant in KO. The methodology offers quantitative and qualitative data as results of the application of six categories of analysis. The most significant scientific production on the Indigenous issue in KO has come from Canada, the United States, and Australia since 1971. In Latin America, publications only began in 2023, particularly in Brazil. We identified two possible paths to improve the subject representation of the area: adaptation of pre-existing schemes or the creation of new knowledge organization systems specialized in Indigenous culture. Cultural hospitality and indigenous warrant are two relevant tools to guide solutions to improve the subject representation of native cultures. Among other conclusions, from the KO, progress was made in the hierarchy of indigenous knowledge, and there was a need for these cultures to impose their ways of categorizing, naming, and relating things. The urgency of promoting academic production on the subject in Latin America is highlighted, considering the historical and contemporary dimension of its great indigenous civilizations throughout its territory.

Received: 11 July 2024; Revised: 27 July 2024; Accepted: 2 August 2024.

Keywords: native peoples; indigenous cultures; knowledge organization systems; cultural warrant; indigenous warrant.

1.0 Introduction

1.1 Native peoples

The situation of many Indigenous cultures in Australia, North, Central, and South America can be described as one of marginalization, subjugation, or minorization (Moulaison-Sandy and Bosaller 2017; Carrón 2019). Although the states attempt various forms of integration or recognition of Indigenous cultures, the truth is that the hegemonic cultures still impose colonialist, racist, and discriminatory traits that affect the social recognition of Indigenous peoples and their cultures. Even today, native cultures continue to be subjugated, ghettoized, or marginalized, with their right to retain their territories often denied or restricted. Long-standing statistical studies provide evidence of the continuity of these marginalization processes in education, rights, and the world of work.

Its members live a parallel existence to that of the descendants of the colonizers, with limited opportunities for real integration. In a book published in 2000, the prolific author of historical and social chronicles, Bryson, described a Saturday street stroll through the city of Alice Springs, located in the arid interior of Australia:

> In the street, there was an overwhelming majority of white people, with a few Aboriginal people also present, walking along the edges of the scene, without disturbing anyone, quietly in the background. The white people did not pay attention to the Aboriginal people, nor did the Aboriginal people pay attention to the white people. It seemed like the two races existed in separate yet parallel worlds. I felt like the only person who considered both groups at the same time (Bryson 2000, 281, translated by the authors).

However, the greatest strength of these Aboriginal communities has been their remarkable resilience. Despite being confined within the heart of dominant cultures and overwhelmed by their hostility or indifference, they have maintained their identity, language, values, beliefs, and interpretations of the world.

Population data concerning indigenous peoples in Latin America, which are essential for any social research, have always been controversial and inconsistent due to various reasons. Since national censuses have been established as a periodic tool for recording population information, one might assume this issue would be coming to a solution. However, registration policies vary from one country to another.

For example, in the United States, any citizen has an immediate answer to the question "What is your race?" because, throughout life, all citizens will have to answer it many times, not only during censuses but also when completing an official form. In contrast, "most Latin Americans alive today have lived their entire lives (...) without having to check a race box on an official form" (Loveman 2014, xi).

The diversity of census registration policies and official records in Latin America, as well as the lack of synchronization in conducting censuses and other population studies, are only two reasons that explain the difficulties in establishing indicators that allow for data comparison and ensuring their reliability.

The continuous and complex processes of integration, acculturation, or mixture of races (Peyser and Chackiel 1994, 93), as well as the difficulties in tracking the composition of migrant populations (a common phenomenon in Latin America), are other factors that distort data. In particular, the migration of indigenous populations assumes various forms (from rural to urban areas, from one country to another, from one region to another), driven by poverty, seasonal work, forced mobilizations, or the impossibility of maintaining their territories (Taylor et al. 2016; Velazco-Ortiz 2023).

A study based on the imprecise and not always compatible data from various national censuses in Latin American countries established 17.4 million indigenous people (Peyser and Chackiel 1994, 100). Other authors double or even triple, suggesting figures as high as 40 to 50 million people depending on the source of information (Del Popolo and Oyarce 2006).

In Brazil, a demographic decline process has been observed. According to Steward (1949), in 1500, there were 1,500,000 indigenous people, a figure that decreased to 500,000 by 1940. More recently, preliminary data collected from the 2022 demographic census found that Brazil has 1,227,642 people identifying as indigenous within Brazilian territory, representing 0.6% of the Brazilian population (IBGE 2022).

The indigenous origins of significant population segments are clear in countries like Mexico, Guatemala, and Bolivia. However, census data assign percentages ranging from 66.2% of the indigenous population in Bolivia to 7.9% in Mexico in 2000 and 2001 (Del Popolo and Oyarce 2006, 41).

According to the Banco Mundial (2015), "while indigenous peoples represent eight percent of the population in the region, they also constitute approximately 14 percent of the poor and 17 percent of the extremely poor in Latin America" (13, translated by the authors). The report adds, "Even today, they face significant challenges in accessing basic services and adopting new technologies, both critical aspects in increasingly globalized societies" (12, translated by the authors).

Beyond various statistical estimates, indigenous peoples have endured centuries of marginalization and segregation, exposed to persecution, wars, and the transmission of dis-

eases – three plagues brought by the colonizers. As a result, they often constitute vulnerable populations economically, culturally, and socially.

In universities, the traditional treatment of indigenous issues – typically rooted in classical Western thought, which can be understood as the perspective of the hegemonic academic culture – has been challenged by a group of intellectuals advocating for new analytical perspectives.

Among the most respected authors are the Brazilian anthropologist and educator Darcy Ribeiro and the Ghanaian-Canadian Georges Dei (2000), who have promoted theoretical frameworks emphasizing the need to decolonize traditional academic thought. They argue that Indigenous knowledge is fundamentally rooted in experience, local perspectives, holistic views, and the spirit of resistance (Dei 2000; Dei and Asgharzadeh 2001).

For its part, the United Nations Educational, Scientific and Cultural Organization (UNESCO) values issues related to indigenous knowledge based on the concept of cultural diversity, which is directly associated with the exercise of human rights and the call to respect the common cultural heritage of humanity.

In Article 4 of its Universal Declaration on Cultural Diversity, UNESCO clearly states, "The defense of cultural diversity is an ethical imperative, inseparable from respect for human dignity. It implies a commitment to human rights and fundamental freedoms, particularly the rights of persons belonging to minorities and those of indigenous peoples. No one may invoke cultural diversity to infringe upon human rights guaranteed by international law, nor to limit their scope" (UNESCO 2001).

1.2 Indigenous Knowledge and Information Science

Within Information Science (IS), the issue of indigenous knowledge is increasingly addressed in texts concerning managing their documentation in libraries and archives and organizing objects preserved in their museums. Libraries, archives, and museums are not isolated entities; instead, they are subordinate to institutions that provide them with economic, material, and human support in exchange for carrying out organized actions to fulfill those institutions' vision, mission, and goals.

In many countries, governments and foundations establish and support institutions to promote the visibility of indigenous cultures by founding and sustaining organizations dedicated to preserving and disseminating the history of native peoples.

Among these, libraries, archives, and museums often serve as primary spaces for gathering, preserving, and disseminating indigenous cultures' material and intellectual evidence. These institutions do not aim to improve the living conditions of indigenous peoples (an objective that other state agencies often attempt with limited and uneven impact) but strive to establish mechanisms for cultural reconstitution and integration. However, they frequently merely represent state forms of social discipline and validation of the idea that indigenous cultures belong to the past.

Littletree et al. (2020) draw on Foucault to highlight his observation that:

> notes the disciplinary function of the state when it comes to the definition of knowledge and the practice of philosophy, and, in turn, how the state utilizes the distinction between kinds of knowledge and forms of inquiry in combination with institutional apparatus such as schools, hospitals, the military, and prisons to discipline – to penalize, order, and conform – its denizens into obedient subjects (Littletree et al. 2020, 412).

In the case of libraries, archives, and museums, disciplining occurs through the knowledge organization of materials and documents under the authority of experts who often do not belong to the indigenous cultures represented. Instead, they express the voice of academic authority, which can inadvertently carry colonialist undertones.

1.3 The Indigenous Issue and Knowledge Organization

The subject representation of indigenous knowledge is a contemporary crossroads since the predominant mentalities of classificationists, classifiers, and indexers are revealed through it. This can consolidate hegemonic visions or propose appropriate alternatives to the cultural particularities of indigenous peoples.

Colonial thought also governs how native cultures are represented in knowledge organization systems (KOS) through ambiguous or inconvenient descriptors or classificatory structures that do not reflect the perspectives and particularities of indigenous knowledge. Critical stances regarding this issue within Knowledge Organization (KO) date back to the 1970s (Yeh 1971; Berman 1978).

In particular, the absence of works by Latin American authors or about Latin American indigenous cultures has been nearly the rule over the past fifty years, even though this region boasts ancient civilizations (such as the Maya, Inca, or Aztec) that have also endured conquest, extermination, and marginalization. These historical processes have been extensively documented in numerous texts on history, sociology, anthropology, and political economy (Galeano 1971).

The concept of warrant allows us to focus on analyzing the terms (and the relationships between terms) that will be selected to constitute the terminological spectrum of op-

tions provided by every KOS for classifying, indexing, and labeling content. As Bullard (2017, 76) points out "warrant is a common thread across a wide variety of systems ranging from traditional library classification to in-application menus and categories for web-based collections [because] all designers of textual organizing schemes must look to some source of terminology".

The concept of cultural warrant was the first within KO to provide theoretical underpinning and support for the need to organize documentation of social groups with particular characteristics integrated into a dominant culture with which they maintain varying degrees of connection – ranging from indifference or peaceful coexistence to resistance and ongoing struggle.

It was Beghtol (2002) who refined and expanded Lee's (1976) original and basic idea, pointing out that cultural warrant "posits that every classification system is based on the assumptions and preoccupations of a certain culture, whether the culture is that of a country, or of some smaller or larger social units (e.g., ethnic group, academic discipline, arts domain, political party, religion and/or language)" (Beghtol 2002, 45).

One form of cultural warrant is, by its nature, the Indigenous warrant. The concept of Indigenous warrant has been developed progressively by Canadian scholar Ann M. Doyle (alone or in collaboration) across various papers committed to creating subject representation spaces that are appropriate for Indigenous knowledge (Doyle 2006, 2013; Webster and Doyle 2008; Doyle et al. 2015; Burns et al. 2017).

Moulaison-Sandy and Bossaller summarize the Doylean concept of the indigenous warrant by stating that "terms and potentially classification structures are derived from the worldview of the indigenous peoples themselves, not from the dominant cultures who write about them or who search for information about them" (Moulaison-Sandy and Bossaller 2017, 133).

Similarly, "Indigenous literary warrant serves as evidence for the classificatory structure and as a source of terminology and is based on indigenous-authored or indigenous-informed literature guided by the primary principle of Indigenous authority" (Doyle et al. 2015, 115). In an attempt to operationalize this warrant, they add that "Indigenous cultural warrant is used, for example, in identifying Indigenous self-representation of names of nations, tribal councils and other forms of governance, as well as contemporary terminology for issues and movements" (115).

2.0 Objectives

Within the broader topic of Indigenous knowledge representation, this study aims to contribute to the theoretical and methodological systematization of the growing body of literature on Indigenous subject representation, particularly Indigenous warrant in KO, from a critical perspective, to promote a Latin American approach. For this reason, beyond the international scope of this study, specific references will be made to the state of the literature in Latin America.

Likewise, the fulfillment of two specific objectives is proposed. Firstly, identify the theoretical and methodological trends presented in the literature for the subject representation based on the Indigenous warrant. Secondly, contribute to teaching the subject representation of indigenous issues in undergraduate and graduate courses in Information Science, particularly in Latin American countries with living indigenous cultures.

3.0 Methodology

The mixed methodology combines qualitative approaches with others that provide quantitative data regarding scientific production on native cultures and subject representation.

To conduct the work, the research question is: What is the state of the art of the debate on Indigenous knowledge representation in KO, especially in relation to the theoretical and methodological consolidation of the so-called indigenous warrant?

The questions associated with the research question and with the objectives mentioned above are: What is the quantitative dimension of the scientific output on the subject representation of native peoples in KO worldwide? How is this production distributed chronologically and geographically? Who are the most productive authors in Indigenous subject representation? What significant theoretical and methodological elements can be highlighted? What are the KOS created to organize the documents and objects of indigenous cultures?

The methodological phases completed the following:

i) Background review in the KO literature: For this purpose, a search was conducted on Google Scholar, updated on June 18, 2024, using the following descriptors: 'Indigenous warrant' (to obtain results in English) and *'garantía indígena*'. This latter expression is written the same way in Spanish and Portuguese, allowing for results in both languages.

   The search was expanded by identifying additional sources mentioned in the bibliographic references of the papers retrieved through Google Scholar. Following the cumulative snowball sampling technique, this second step was implemented to add documents that did not explicitly contain the term 'Indigenous warrant' but were related to the studied topic. The gathered items were refined by excluding those that were not relevant

to the research. It was determined that only those papers which significantly offered content related to the subject representation of Indigenous knowledge would be included in the corpus.

ii) Formation of the corpus: The results obtained from the Google Scholar search were as follows: 'Indigenous warrant' yielded 31 results, and 'garantía indígena' yielded 10 results, totalizing 41 initial papers. This list was refined by excluding irrelevant content for the research objective and removing duplicate results. Thus, the initial corpus was reduced to 24 papers. Next, an analysis of the references found in the 24 papers was conducted. Through this process, an additional 39 references were obtained, resulting in a final corpus of 63 papers, all of which are cited in Table 1. Full texts were accessible in 85% of the cases, while abstracts and other information, such as reviews or compilations, were available in the remaining 15%. The references for each paper included in the corpus are interspersed within the references of this article.
Thus, the corpus was integrated with papers that include content related to KOS and/or classification, indexing, terminology, or the language specific to Indigenous peoples in catalogs or databases, emphasizing those focused on indigenous warrant.
With this delimitation, valuable documents unrelated to the scope of this work were excluded, such as those related to library services, the role of archives and archivists, historical aspects, and technological aspects related to the management of documentation of Indigenous peoples.

iii) Categories of analysis: Once the corpus had been constituted, the following categories of analysis were established to conduct both a formal analysis and a critical reading of the corpus:

- geographical reference of the works;
- chronological reference of the works;
- production by authors;
- theoretical postulates raised in knowledge organization;
- an inventory of KOS was designed to organize Indigenous knowledge.
- techniques and methodologies for the application of Indigenous warrant.

iv) Finally, the results were organized as presented in the next chapter.

4.0 Results

4.1 Geographical reference of the works on Indigenous peoples

The 63 papers comprising the corpus were categorized by country, considering the country of institutional affiliation of the author or the first author in the case of co-authored papers.

Table 1 displays the results indicating the country, citations, and number of citations.

4.2 Chronological reference of the works

In Table 2, the 63 papers are shown as being distributed for decades from 1970 onwards, since the earliest identified contribution is dated 1971.

4.3 Production by authors

In Table 3, authors are presented in descending order based on their academic production. Both individual authorships and co-authorships are counted equally. Specific data is provided only for authors with two or more published papers.

In addition to the nine authors with two or more authorships, we must add more than 80 authors with only one publication to date, which brings the number of authors who have produced material on the subject closer to one hundred from 1971 to this date.

4.4 Theoretical postulates raised in knowledge organization

Due to the critical approach made on the information collected on the theoretical aspects involved in the organization of Indigenous knowledge, both the results and their discussion are developed in section 5.4.

4.5 Inventory of KOS designed to organize Indigenous knowledge.

The growing awareness of the need to indigenize KO has been expressed in the publication of numerous Indigenous knowledge organization systems (IKOS) in recent years. These systems have been created to organize documentary and material collections of indigenous cultures in various parts of the world. In general terms, they share the same theoretical postulates, although each IKOS has been built based on its own design decisions. They are only applied in their countries of origin because they are intended to address the subject representation of Indigenous knowledge specific to their ethnic groups.

Country	Citations	Number of citations
United States	Beall 2006; Berman 1978, 1995; Buente et al. 2020; Burns et al. 2017; Camacho 2023;Campbell et al 2022; Carrón 2019; Duarte and Belarde-Lewis 2015; Frosio 1971; Gosart 2021; Green 2015; Hajibayova and Buente 2017; Kam 2007; Levinson 2023; Littletree and Metoyer 2015; Littletree 2019; Littletree et al. 2020; Moulaison-Sandy and Bossaller 2017; Nyitray and Reijerkerk 2021; Pettitt and Elzi 2023; Tomren 2004; Webster and Doyle; 2008; Yeh 1971; Young and Doolittle 1994	25
Canada	Boisvert 2023; Bone and Lougheed 2018; Bosum and Dunne 2017; Cherry and Mukunda 2015; Chester 2006; Doyle 2006, 2013; Doyle et al. 2015; Dudley 2017; Farnel 2021; Farnel et al. 2016; Gilman 2006; Godbold 2009; Knight 2019; Lee 2011; Lee et al. 2021; MacDonell et al. 2003; Olson 1999; Swanson 2015	19
New Zealand	Bardenheier et al. 2015; Bryant 2015; East 2008; Lilley 2015; Simpson 2005; Szekely 1997	6
Australia	Moorcroft 1993, 1994, 1997; Nakata and Langton 2005; Thorpe and Galassi 2014	5
Brazil	Albuquerque and Moraes 2023; Gracioso et al. 2023; Moraes 2023; Silva 2023	4
Portugal	Simões 2023	1
Thailand	Chongchorhor and Kabmala 2022	1
Uruguay	Barité and Moutinho 2023	1
Zimbabwe	Maware 2012	1
Total		63

Table 1. Corpus by countries of institutional affiliation of the authors

Decade	Works	%	Progression %
1970-1979	3	4,8	4,8
1980-1989	0	0	4,8
1990-2000	7	11,1	15,9
2001-2010	12	19	34,9
2011-2020	25	39,7	74,6
2021-2023	16	25,4	100
Total	63	100	100

Table 2. Chronological distribution of works by decades

Authors	Number of works	Years	Country
Doyle, Ann Mary	5	2006, 2008, 2013, 2015, 2017	Canada
Littletree, Sandra	3	2015, 2019, 2020	United States
Moorcroft, Heather	3	1993, 1994, 1997	Australia
Belarde-Lewis, Miranda	2	2015, 2020	United States
Berman, Sanford	2	1978, 1995	United States
Buente, Wayne	2	2017, 2020	United States
Duarte, Marisa	2	2015, 2020	United States
Dupont, Sarah	2	2015, 2021	Canada
Farnel, Sharon	2	2016, 2021	Canada

Table 3. Works distribution by authors

Table 4 lists the most recognized IKOS in the literature. Some ongoing projects could be added to this list, such as the ontology that is being designed to organize the knowledge of the Thai ethnic group in Thailand (Chongchorhor and Kabmala 2022).

4.6 Techniques and methodologies for the application of Indigenous warrant.

Those who promote the creation of new KOS intended to represent indigenous knowledge do not, in general, propose

System	Type	Country	Site or literature
Pathways: Gateway to the AIATSIS Thesauri (2010)	Thesaurus	Australia	https://www1.aiatsis.gov.au/
Brian Deer Classification System (BDCS)	Classification system	Canada	Carron 2019
Xwi7xwa Classification Scheme	Classification system	Canada	https://xwi7xwa.library.ubc.ca/collections/indigenous-knowledge-organization/
Mashantucket Pequot Thesaurus	Thesaurus	United States	Littletree and Metoyer 2015
Māori Subject Headings	Subject headings	New Zealand	Bardenheier et al 2015.; Lilley 2015

Table 4. List of IKOS

methodological innovations in the design of the schemes. Rather than new techniques or methodologies, the emphasis is placed on prioritizing the language of native peoples, and on the distribution of vocabulary in classes that respect the worldviews of these communities.

However, those who focus on the adaptation of pre-existing schemes, proposed, in essence, modalities of application of the principle of cultural hospitality. As Choi et al. (2022) point out, cultural hospitality constitutes "an approach to improve information systems by providing ethical resource descriptions and access" (554). In this way, "cultural hospitality refers to the ability of a system to connect existing knowledge with perspectives, expectations, and assumptions from different cultures and users" (554).

Another element to consider in the framework of cultural hospitality is that the choice of descriptors in Aboriginal themes must respect the concepts as they are constructed and named in their culture of origin (Farnel 2010; Moulaison-Sandy and Bossaller 2017).

5.0 Discussion

5.1 Geographical reference of the works on Indigenous peoples

As can be seen, the authors come from universities or institutions in only nine countries: Australia, Brazil, Canada, New Zealand, Portugal, Thailand, the United States, Uruguay, and Zimbabwe.

Furthermore, there is a clear predominance of production from the United States and Canada, as these two countries account for 44 papers, nearly 72% of the total.

Only five works come from Latin American authors (4 from Brazil and 1 from Uruguay, in the latter case with Brazilian co-authorship).

The five works were published in 2023, expressing the absolute novelty of treating indigenous issues in KO ¬literature in the Latin American region. This is particularly suggestive if we consider that most Latin American countries have ancient indigenous populations with similar realities to those found with the native peoples of North America, Australia, and New Zealand.

The study by Gracioso et al. (2023) deals with the challenges of KO for the subject representation of knowledge about Indigenous peoples in information systems. The authors point out that the growing participation of Indigenous people in Brazilian universities, one of the main achievements derived from the struggle of Indigenous peoples for their rights, has impacted the production of research, requiring the establishment of indexing policies of institutional repositories, which allow keywords to be offered and used in the original languages of the people who produce the research,

The work of Moraes (2023) seeks to build a terminological instrument that responds to both the principle of literary warrant and the decolonial perspective and takes as reference a glossary by Cavalcanti Proença, based on the work Macunaíma by Mário de Andrade. Moraes identifies 2,112 terms and synonyms (generally indigenous voices), each followed by a definition, which gives a dimension to the richness of the Brazilian vocabulary from its roots. The study concludes that even facing representation problems similar to those of interdisciplinary spaces, Macunaíma's Decolonial Glossary can contribute to documenting the National Inventory of Linguistic Diversity, instituted in 2010, to safeguard Brazilian indigenous languages. The resulting vocabulary can be used for different subject representation operations if necessary.

In the case of Silva (2023), it is a master thesis defended at the University of San Carlos, Brazil, in which the author describes the initiatives to create and maintain lists of subject headings and thesauri in light of the justice and social equity, based on the literature on the sociocultural dimension of KO. The study identifies discussions about the subject representation of different social groups in lists of subject headings and thesauri, including Indigenous communities, to which she dedicates two segments of her thesis. To obtain inclusive subject representations for minority social groups, Silva proposes implementing local modifications of the KOS, based on the language of the local culture.

For their part, Albuquerque and Moraes (2023) relate knowledge to different cultures, which are made up of historical subjects and discursive communities, in the scope of more general social processes. The authors critically look at how the specificities of Indigenous cultures are treated in SI in Brazil while identifying the epistemological, theoretical, and methodological traditions of KO in articles from Brazilian journals, which can contribute to improving the Indigenous subject representation and user knowledge.

Finally, Barité and Moutinho (2023) carry out a critical review of the existing literature on forms of Indigenous warrant in KO, based on four categories of analysis: identification of criteria for the organization of knowledge specific to Indigenous cultures; characterization of the Indigenous warrant as a kind of cultural warrant; identification of the institutionalization processes of documentation and objects specific to Indigenous cultures; and, identification and description of methodologies for the application of the Indigenous warrant. It is concluded that there is a sound theoretical and methodological basis to consolidate the concept and application of the Indigenous warrant. Likewise, given the almost non-existent Latin American literature on the subject, the urgency of promoting academic production on the subject representation of indigenous cultures in Latin America is mentioned.

5.2 Chronological reference of the works

The production of subject representation and native people has increased significantly in the last twenty-three years, and the progression does not appear to be stopping.

If the data in Tables 1 and 2 are related, it can be verified that most works published between 2001 and 2020 belong to Canadian and American authors. However, from 2021 to date, the production of these two countries has slowed down, while the first Latin American publications (Brazil and Uruguay) emerged.

5.3 Production by authors

The nine most active authors belong to only three countries: three from Canada, five from the United States, and one from Australia.

The most frequent co-authorship occurs in the two papers by Duarte and Belarde-Lewis. Ann Mary Doyle from Canada stands out prominently for the depth and originality of her approaches and her ability to collaborate on papers.

Another noteworthy point is that, apart from Doyle and Berman, the most productive authors have published their works within a span of five years.

5.4 Theoretical postulates raised in KO

As Agrawal (2002, 87) points out, "It would be fair to claim that the contemporary attention to indigenous knowledge is in no small measure as a result of its successfully posited connection with development and environmental conservation". This reference is significant because it tacitly expresses the need to understand the life of native peoples and their ways of interpreting reality within an ecosystem where nature and humanity coexist under different rules than those of the so-called Western civilization. Those responsible for the KOS who have the challenge of integrating Indigenous knowledge into their schemes have to make an effort to identify, first of all, the cultural keys of the native peoples, their ways of governing themselves, their values, their rituals, and their principles.

The most significant problems that have arisen in the treatment of Indigenous knowledge by those responsible for the main universal classification systems (Universal Decimal Classification-UDC, Dewey Decimal Classification-DDC, Library of Congress System-LCC and Library of Congress Subject Headings-LCSH) can be summarized in one sentence: "the literature documents that the mainstream systems tend to marginalize, omit or misrepresent Indigenous topics. These types of inaccuracies can occur through historicization, lack of specificity, lack of relevance, lack of recognition of sovereign nations, and the omission of the historical realities of colonization" (Burns et al. 2017, 2040). By opting for classical organizations of knowledge based on Western thought, universal KOS have difficulty incorporating what is different or diverse. As Szostak (2014, 160) points out, the "existing classifications privilege certain ways of looking at the world while obscuring others".

Among different reasons that explain the dissatisfaction of professionals and users of libraries, archives, and museums who used these systems over time, it is noted that universal systems have remained hostage to their ambition for universality and international reach since the result is that their schemes end up validating the hegemonic conceptions that accompany the development of global projects and make cultures invisible or marginalized.

The organization of Indigenous knowledge involves incorporating new theoretical postulates, which also lead to reviewing the methods and techniques used for the description, classification, and indexing of resources of Indigenous peoples, as well as accepting the idea of creating specific schemes or systems.

What new premises do the authors identify to promote alternative conceptions to the hegemonic ones?

First, Indigenous cultures should be placed on an equal footing, prioritizing their culture and traditions and seeking to understand and overcome the reasons why Indigenous knowledge tends to be shown as inferior (Doyle et al. 2015).

At least two of the mechanisms of marginalization of indigenous cultures are mentioned: firstly, the idea, quite deeply rooted in contemporary urban societies, that indigenous cultures are part of the past. Authors such as Doyle, Littletree, and Farnel propose to follow another path, assuming that the original cultures were not only before the current hegemonic culture but that their cultural heritage is still alive and current because those Indigenous communities still exist.

Secondly, recognize that native peoples have other forms of knowledge transmission, particularly oral tradition (Moulaison-Sandy and Bossaller 2017; Carrón 2019) and the regular development of generally sacred or mystical rituals (Camacho 2023). Indeed, oral recording is always possible: it can be recorded, filmed, or taken to printed sources, interviews, or forms of artistic and cultural expression. Even so, this does not always constitute a priority or an internal need for the members of a native culture.

Thirdly, the requirement is to understand the points of view of the people whose ideas are represented (Green 2015; Littletree and Metoyer 2015; Moulaison-Sandy and Bossaller 2017). As Rosztak (2014, 161) points out "It is generally difficult to identify the dominant perspective of a particular social group. [It is much easier to identify the dominant perspective of an academic discipline.]" One way to overcome this obstacle is to ensure the participation of members of the Indigenous community in decision-making as members or consultants of teams of classificationists.

In this sense, Farnel (2021) points out that a criterion for establishing culturally sensitive metadata must necessarily incorporate the knowledge structures used by culture members (Farnel 2021, 8). Thus, it implies complying with the steps that lead to using the minority culture's ways of thinking, organizing, and designating, avoiding any linguistic-ideological bias that may come from the hegemonic or dominant culture. To do so, it took three concepts as its foundation: i) the anticolonial theory that "emphasizes the multiplicity of local Indigenous knowledge, and asserts their ability to resist colonial power structures and to go beyond dismantling colonial structures by building new and better structures based on that knowledge (Farnel 2021, 3); ii) the theory of fluid ontologies promoted by Srinivasan (2002, 2007), to establish flexible knowledge structures that consider the interests of communities with their own culture; and, iii) the sociolinguistic theory of language codes (Farnel 2021).

As Wise and Kostelecky (2018) concluded, collaboration with members of one Indigenous community (in their case, the Zuni people) dramatically improved item description, collection discoverability, and collection interactivity.

With the consultation of Indigenous opinion, for example, the Indigenous names of places, rituals, music, plants, tools, and any other object typical of that culture could be incorporated into the schemes instead of names translated into English or to another reference language, or to generic names that do not faithfully reflect the specificity and diversity of Indigenous knowledge,

Regarding theoretical-methodological trends in the corpus that provides the basis for this work, two possible paths have been identified to date to improve the subject representation of the documentation and objects that constitute the heritage of Indigenous culture: i) the adaptation of already existing classification schemes, with the focus on the visibility and organization of Indigenous knowledge; and, ii) the creation of KOS intended to represent indigenous knowledge exclusively, with the focus on indigenizing KO, which is also known as Indigenous Knowledge Organization or IKO (Doyle 2006; Doyle et al. 2015). Carrón (2019) makes a good summary of the current state of this dilemma and offers a broad description of the traditional procedures for inscribing indigenous topics in systems such as DDC and LCSH, while exploring the creation of alternative classification standards metadata schemes and new digital platforms and tools to facilitate discovering information for and about Indigenous people.

Adapting classification schemes offers partial solutions to resolve the absence, insufficiency, or inconsistency of matters with literary warrant (that is, with sufficient supporting documentation), which are important to many users. Those who have chosen this traditional path have offered specific techniques for inserting indigenous topics mainly in LCSH (Beall 2006; Lee 2011; Bone and Lougheed 2018; Campbell et al. 2022; Pettitt and Elzi 2023), but also to LCC (Yeh 1971) or UDC (Simões 2023). These adaptations can be seen as contributions to be incorporated by the teams responsible for the KOS in future editions or as unauthorized and local solutions to resolve the relationship between the documentary collections and materials of an ethnic group and its users.

Indigenizing KO, on the contrary, entails assuming a radical change since it requires building a new epistemology. It implies the proposal of new forms of subject representation based on new theoretical bases. Already in 2006, Doyle (110) advocated the need to 'indigenize' IS by developing theoretical and conceptual frameworks that would allow professional tools to be adapted to the needs and purposes of indigenous cultures. Doyle proposed indigenizing the discipline because IS did not treat these cultures from their perspectives. As he points out a decade later in a co-authored work, the forms of organization of Indigenous materials and documents, as well as the terminology used for subject representation, came from the visions and values of "newcomers to First Nations territories including early anthropologists, missionaries, government agents, and travelers, and not Indigenous perspectives or values" (Doyle et al. 2015, 111).

In this second document, Doyle et al. (2015) raise the need to 'indigenize' KO, analyzing "possible intersections between Indigenous frameworks and the information professions" (115). They highlight the bi-directionality of these processes: "We seek these intersections in order to explore ways in which KO might serve Indigenous interests, and ways in which to indigenize the discipline of KO itself; "this is both a critical and constructive undertaking" (115).

Indigenizing KO means not only considering the material and immaterial evidence of Indigenous culture as expressions of the heyday of peoples who were later absorbed by the advances of 'civilization', but also as a proof of the vitality of cultures that come from the depths of the history of our countries, and have demonstrated extraordinary resilience, and an extreme attachment to the defence of their identity and traditions. In the words of Littletree et al. (2020), this implies practicing an epistemological intervention where Indigenous artifacts, relics, and documents are not seen as individual objects but as integrating elements of a tradition and a cultural construction. In this sense, the authors propose relationality as the organizing principle of this cultural construction for the identification, discernment, creation, and continuation of Indigenous knowledge systems (413). The authors argue that

> to understand IKO – that is, the methodologies and means by which Native and Indigenous peoples create protocols to cohere, name, articulate, collate, and make accessible objects that indicate Indigenous knowledge – requires that practitioners of KO appreciate the colonial history of KO. Furthermore, it requires that KO practitioners recognize that the work of IKO is fundamentally a practice of liberation (Littletree et al. 2020, 413).

In any case, both currents (that of local adaptation and the creation of new KOS) agree on the insufficiency of traditional classification schemes to offer adequate and ethically irreproachable subject representations of indigenous concepts.

A milestone in the theoretical discussion on the subject representation of indigenous knowledge is the publication of an issue of the journal Cataloging & Classification Quarterly (number 5/6 of volume 53, corresponding to the year 2015), coordinated by Ann M. Doyle and Cheryl Metoyer, dedicated to the organization of Indigenous knowledge.

5.5 Inventory of KOS designed to organize Indigenous knowledge

Pathways is an initiative of the Australian Institute of Aboriginal and Torres Strait Islander Studies. It integrates three thesauri (for place names, languages and peoples, and disciplines) and is an extension of the original publication The Aboriginal and Torres Strait Islander Thesaurus, developed by Heather Moorcroft and Alana Garwood and published by the National Library of Australia in 1997 (Lee 2011)

The Brian Deer Classification System (BDCS) is a system created in the 1970s to accurately reflect Indigenous ways of knowing from a First Nations perspective through the representation of their histories, words, and worldviews (Carrón 2019; Duarte and Belarde-Lewis 2020).

Xwi7xwa Classification Scheme is based on the BDCS and was adapted for use in several Canadian libraries. In 2004, the Xwi7xwa Library of British Columbia "applied to the Library of Congress MARC Standards Office to legitimize the scheme on an international level, [and] in 2005, the request was granted, and the new scheme (officially termed as the First Nations House of Learning (FNHL) Subject Headings was officially authorized as a thesaurus "which could then be fully indexed in the authorized subject headings MARC field (650)" (Doyle et al. 2015, 113).

The Mashantucket Pequot Thesaurus is a product of the American Indian Terminology Project, which Sandra Littletree and Cheryl Metoyer have led from the University of Washington. According to those responsible, this unpublished thesaurus "is designed to be user-centered and to reflect the information seeking behavior of Native and non-Native scholars and researchers who conduct research on American Indians and as a controlled vocabulary; the primary goal of the Thesaurus is to inform Library of Congress Subject Headings" (Littletree and Metoyer 2015, 641).

For its part, Māori Subject Headings (MSH) is a structured list of descriptors related to the Māori culture of New Zealand. The list was created in 2006, under the responsibility of the National Library of New Zealand, to provide terms familiar to Māori people and arranged in a hierarchy that reflects the Māori worldview.

5.6 Techniques and methodologies for the application of the Indigenous warrant

As mentioned above, it can be accepted that Indigenous warrant is a variety of cultural warrant. This statement derives from recognizing that using a consistent battery of theories and methodologies typical of cultural warrant in the more restricted indigenous knowledge organization field is possible.

Within the strictly theoretical approaches, a work by Olson (1999) stands out, in which the cultural construction of the classifications made according to Western thought is explored (taking Aristotle, Durkheim, and Foucault as references), and the possibility of conceiving alternative classification forms. To demonstrate that the latter is possible, Olson analyzes how Indigenous cultures use the criteria of exclusivity, teleology, and hierarchy and forces us to think

about classification schemes constructed from other perspectives.

In this regard, it is worth mentioning at least two antecedents of techniques and methods that can be associated with cultural warrant. Barité and Rauch (2020) have proposed methods common to some social sciences and humanities, such as content analysis, terminological analysis, discourse analysis, and, already within KO, the techniques and methods of domain analysis.

For their part, Olson and Ward (1998) suggested creating paradoxical spaces to insert gender terminology in the DDC tables. This solution is extensible to any other situation in which it is necessary to introduce culturally oriented terms in the schemes of a KOS. Implementing paradoxical spaces involves creating a new term that does not exist in a KOS, which is the opposite of one accepted by conventional or traditional decisions.

The methods and techniques of Indigenous warrant have the same application difficulties as all qualitative modalities. The results cannot be measured with indisputable fidelity, and they may involve biases or deviations inherent to the mentalities of the analysts or the interpretation criteria.

For example, it has been said that the language of indigenous cultures should be privileged, but how is it possible to determine this? Even the consultants from the indigenous communities involved may have differences regarding the choice of candidate descriptors.

In any case, generic guidelines can be provided, considering the problematic areas that the Indigenous warrant application processes must resolve: the transition from universal KOS with inadequate subject representation to KOS with original Indigenous schemes; the replacement of representations imposed by hegemonic cultures with forms of subject representation typical of Indigenous cultures; the elimination of terms that connote forms of marginalization and the introduction of terminology that empowers and prioritizes Indigenous conceptions; the generation of strategies that show the vitality and validity of the cultures of native peoples; the identification of omissions, lack of specificity and inaccuracies, to replace them with inclusive, specific and precise. terms.

As Figure 1 shows, these guidelines can be used as generic parameters for applying the Indigenous warrant and its evaluation.

6.0 Concluding remarks

Native peoples constitute significant minorities in much of South America, Central America and the Caribbean, North

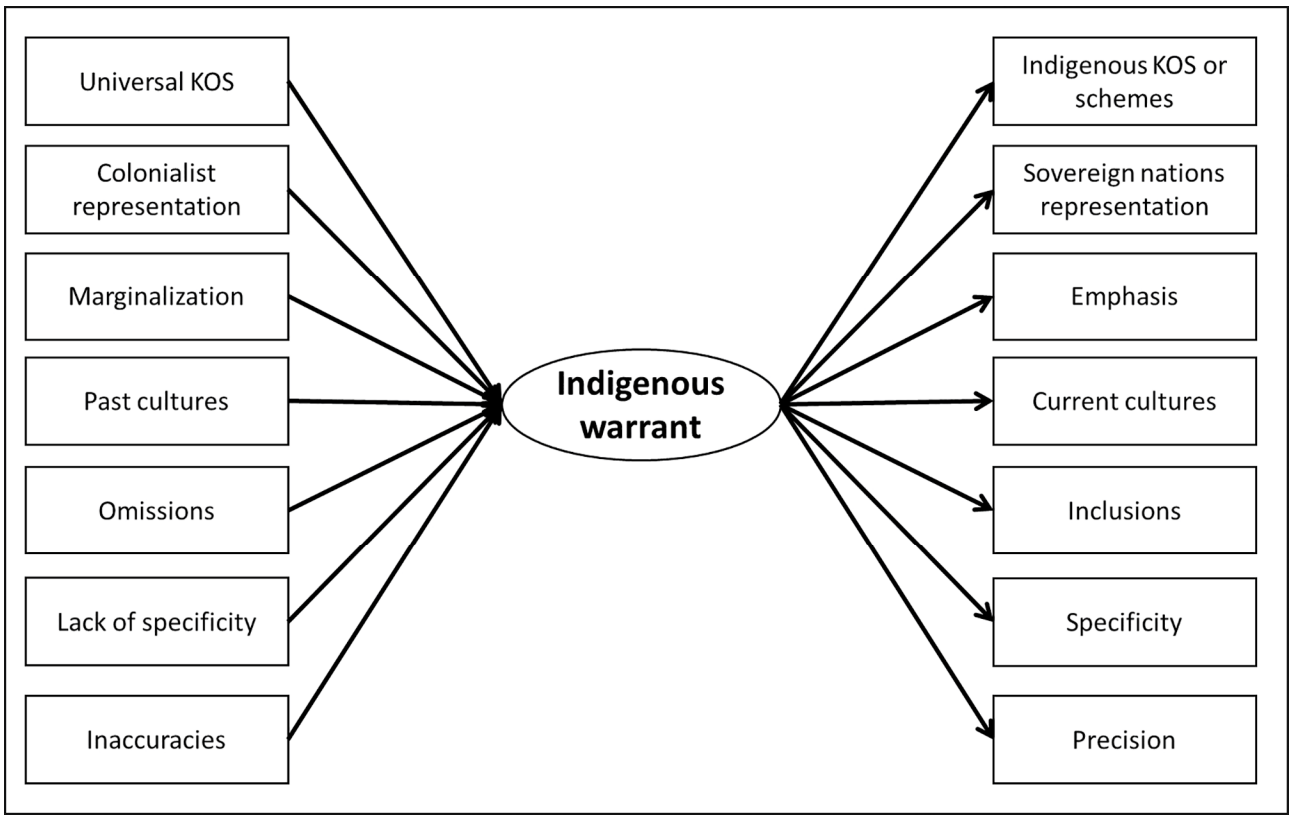


Figure 1. Guidelines for applying the Indigenous warrant.

America, Africa, and Oceania. Each indigenous culture constitutes a sociocultural ecosystem, surrounded to a certain extent by a hostile environment or, at least, indifferent to its fate and destiny. People with common rituals, beliefs, and values interact in this relatively closed ecosystem. Among other peculiarities, Indigenous knowledge is characterized by a long tradition of oral transmission. This is why it is more common to find documentation on indigenous cultures generated by the conquerors and, more recently, by social researchers trained within current Western thought, than by the same native communities. Only in the last 30 years have professionals dedicated to collecting, preserving, and organizing this documentation and the material evidence of native cultures in archives, libraries, and museums. Thus, they have realized that a new theoretical framework was needed to guide their practices.

From the KO, progress was made in the hierarchy of Indigenous knowledge, and there was a need for these cultures to impose their ways of categorizing, naming, and relating when specifying universal classification schemes or creating KOS intended for Indigenous collections. In a significant number, the latter can already be offered to integrate or complement traditional KOS or be used independently.

The cultural hospitality principle, born alongside the terminological selection criterion called cultural warrant, constitutes a good tool for developing reliable methodologies for subject representation. As the Indigenous warrant is a variety of cultural warrant, it is possible to use their consistent battery of theories and methodologies to guide the application of the Indigenous warrant.

The research demonstrates that the Latin American approaches to the Indigenous issue from KO are highly new since the pioneering contributions date back to 2023, which marks the dimension of oblivion and indifference of the Latin American KO in a continent with ancient indigenous cultures.

However, countries such as the United States, Canada, and Australia have developed a body of literature ensuring that future Latin American researchers have a significant epistemological and conceptual basis to frame the KO of their indigenous cultures.

Given the universality and similarity of the processes of marginalization and institutional and cultural recomposition of indigenous heritage between these countries and those of Latin American countries, they can serve as a model for finding solutions to indigenous KO in Latin America.

The urgency of promoting academic production on the subject in Latin America is highlighted, considering the historical and contemporary dimension of its great indigenous civilizations throughout its territory.

Through the Indigenous warrant used by participatory forms of management and curation of libraries, archives, and museums, it is possible to strengthen respect for universal values expressed in local cultures and increasingly associated with citizenship construction, social inclusion, and respect for alternative cultures.

Note

This work is included in the research activities financed by the Scientific Research Commission (CSIC) of the University of the Republic of Uruguay.

References

Agrawal, Arun. 2002. “Indigenous Knowledge and the Politics of Classification”. *International Social Science Journal* 54, no. 173: 287-97. https://doi.org/10.1111/1468-2451.00382

Albuquerque, Ana Cristina and Marcos A. Moraes. 2023. “A organização e representação do conhecimento indígena a partir de um estudo nas publicações da área de Ciência da Informação no Brasil”. In *Memoria del Encuentro de Directores y de Docentes de Escuelas de Bibliotecología y Ciencia de la Información del MERCOSUR*. https://encuentro-mercosur.fic.edu.uy/index.php/encuentro-mercosur/article/view/60

Banco Mundial. 2015. *Latinoamérica Indígena en el Siglo XXI*. Washington, D.C.: Banco Mundial.

Barité, Mario and Mirtha Rauch. 2020. “Cultural Warrant: Old and New Sights from Knowledge Organization”. In *Knowledge Organization at the Interface: Proceedings of the Sixteenth International ISKO Conference, 2020, Aalborg, DK*, edited by Marianne Lykke, Tanja Svarre, Mette Skov and Daniel Martínez-Ávila. Advances in Knowledge Organization 17. Würzburg: Ergon Verlag, 31-40.

Barité, Mario and Sonia Matos Moutinho. 2023. “La garantía indígena: aportes para su aplicación al patrimonio cultural de los pueblos aborígenes latinoamericanos”. In *Organização e Representação do Conhecimento em Diferentes Contextos: Desafios e Perspectivas na era da Datificação*, edited by Natália Bolfarini Tognoli, Ana Cristina de Albuquerque and Brígida Maria Nogueira Cervantes, 650-60. Londrina: ISKO-Brasil.

Beall, Jeffrey. 2006. “Ethnic groups and Library of Congress subject headings”. *Colorado Libraries* 32, no. 4: 37-44.

Beghtol, Clare. 2002. “A Proposed Ethical Warrant for Global Knowledge Representation and Organization Systems”. *Journal of Documentation* 58, no. 5: 507-32. https://doi.org/10.1108/00220410210441

Belarde-Lewis, M. 2013. “From Six Directions: Documenting and Protecting Zuni Knowledge in Multiple Environments”. PhD diss., University of Washington.

Berman, Sanford. 1978. "Stop Playing Hide-and- Seek with Ethnic Materials". *Wilson library bulletin* 52: 691-719.

Berman, Sanford. 1995. "When the Subject Is Indians". *American Indian Libraries Newsletter* 18, no. 2.

Bone, Christine and Brett Lougheed. 2018. "Library of Congress Subject Headings Related to Indigenous Peoples: Changing LCSH for Use in a Canadian Archival Context". *Cataloging & Classification Quarterly* 56, no. 1: 83-95. https://doi.org/10.1080/01639374.2017.1382641

Bosum, Annie and Ashley Dunne. 2017. "Implementing the Brian Deer Classification System for Aanischaaukamikw Cree Cultural Institute". *Collection Management* 42, no. 3-4: 280-93. https://doi.org/10.1080/01462679.2017.1340858

Bryson, Bill. 2000. *En las Antípodas*. Barcelona: RBA.

Buente, Wayne, Chad Kālepa Baybayan, Lala Hajibayova, Mallory McCorkhill and Roman Panchyshyn. 2020. "Exploring the Renaissance of Wayfinding and Voyaging through the Lens of Knowledge Representation, Organization and Discovery Systems". *Journal of Documentation* 76, no. 6: 1279-93. https://doi.org/10.1108/JD-10-2019-0212

Bullard, Julia 2017. "Warrant as a Means to Study Classification System Design". *Journal of Documentation* 73, no. 1: 75-90. https://doi.org/10.1108/JD-06-2016-0074

Burns, Kathleen, Ann Mary Doyle, Gene Joseph and Allison Krebs. 2017. "Indigenous Librarianship". In *Encyclopedia of Library and Information Science,* edited by John D. McDonald and Michael Levine-Clark, 2031-47. 4th ed. Boca Raton: Francis & Taylor.

Camacho, J. 2023. "Decolonizing Libraries to Advance Ethical Resource Classification and Description for Traditional Knowledge". Master's thesis, University of California

Carrón, Tara. 2019. "Integrating Indigenous Knowledge in Non-Indigenous Knowledge Systems and Institutions". *John Hopkins University: Museum Studies Digital Curation Certificate Program*. https://jscholarship.library.jhu.edu/handle/1774.2/62121

Cherry, Alissa and Keshav. Mukunda. 2015. "A Case Study in Indigenous Classification: Revisiting and Reviving the Brian Deer Scheme". *Cataloging & Classification Quarterly* 53, no. 5-6: 548-67. https://doi.org/10.1080/01639374.2015.1008717

Chester, G. G. 2006. "Proposed Tribal College Cataloging Systems: From Isolation to Association". PhD diss., University of Minnesota.

Choi, Inkyung, Wan-Chen Lee, Ying-Hsang Liu, Hsinliang Chen, Douglas W. Oard and Chi Young Oh. 2022. "Cross-cultural Information Access". *Proceedings of the Association for Information Science and Technology* 59, no. 1: 551-54. https://doi.org/10.1002/pra2.624

Chongchorhor, Chokthumrong. and Malee Kabmala, 2022. "The Development of an Ontology for Thai's Indigenous Rice Knowledge in Thailand". *Journal of Library and Information Studies* 20, no. 2: 55-74. https://doi.org/10.6182/jlis.202212_20(2).055

Dei, George J. Sefa. 2000. "Rethinking the Role of Indigenous Knowledges in the Academy". *International Journal of Inclusive Education* 4, no. 2: 111-32. https://doi.org/10.1080/136031100284849

Dei, George J. Sefa and Alireza Asgharzadeh. 2001. "The Power of Social Theory: The Anti-Colonial Discursive Framework". *The Journal of Educational Thought (JET) / Revue de la Pensée Éducative* 35, no. 3: 297-323. http://www.jstor.org/stable/23767242

Del Popolo, Fabiana and Ana María Oyarce. 2006. "Población indígena de América Latina: perfil sociodemográfico en el marco de la conferencia internacional sobre la población y el desarrollo y de las metas del milenio". In *Pueblos Indígenas y Afrodescendientes de América Latina y el Caribe: Información Sociodemográfica para Políticas y Programas: Documento de Proyecto*, edited by Comisión Económica para América Latina y el Caribe (CEPAL), 35-62. Santiago de Chile: ONU.

Doyle, Ann Mary. 2006. "Naming and Reclaiming Indigenous Knowledges in Public Institutions: Intersections of Landscapes and Experience in Knowledge Organization for a Global Learning Society". In *Knowledge Organization for a Global Learning Society: Proceedings of the Ninth International ISKO Conference 4-7 July 2006 Vienna, Austria*, edited by Gerhard Budin, Christian Swertz and Konstantin Mitgutsch. Advances in Knowledge Organization 10. Würzburg: Ergon Verlag, 436-42.

Doyle, A. M. 2013. "Naming, Claiming, and (Re)creating: Indigenous Knowledge Organization at the Cultural Interface". PhD diss., University of British Columbia.

Doyle, Ann Mary, Kimberley Lawson and Sarah Dupont. 2015. "Indigenization of Knowledge Organization at the Xwi7xwa Library". *Journal of Library and Information Studies* 13, no. 2: 107-34. https://doi.org/10.14288/1.0103204

Dudley, Michael Q. 2017. "A library matter of genocide: the Library of Congress and the Historiography of the Native American Holocaust". *International Indigenous Policy Journal* 8, no. 2. https://doi.org/10.18584/iipj.2017.8.2.9

Duarte, Marisa Elena and Miranda Belarde-Lewis. 2015. "Imagining: Creating Spaces for Indigenous Ontologies". *Cataloging & Classification Quarterly* 53, no. 5-6: 677-702. https://doi.org/10.1080/01639374.2015.1018396

East, Robyn. 2008. "Indigenous Subject Access: Ngā Ūpoko Tukutuku the Māori Subject Headings Thesaurus". Paper presented at the Australian and New Zealand Theoretical Library Association (ANZTLA) conference, Wellington, New Zealand, July 2008.

Farnel, Sharon. 2021. "Knowledge Organization in Support of an Indigenous Worldview: A Community Driven Metadata Framework for the Inuvialuit Digital Library". *Proceedings from North American Symposium on Knowledge Organization* 8: 1-17. https://doi.org/10.7152/nasko.v8i1.15859

Farnel, Sharon, Ali Shiri, Dinesh Rathi, Cathy Cockney, Sandy Campbell, and Robyn Stobbs. 2016. "Of places and names: working with northern Canadian communities to enhance subject access to digital resources". In *Connections. Collaboration. Community. 82nd IFLA World Library and Information Congress, Columbus, OH, 13-19 August 2016*. https://library.ifla.org/id/eprint/1326/1/151-farnel-en.pdf

Frosio, Eugene T. 1971. "Comments on the Thomas Yen-Ran Yeh Proposals". *Library Resources and Technical Services* 15, no. 2: 128–31.

Galeano, Eduardo. 1971. *Las venas abiertas de América Latina*. México City: Siglo XXI Editores.

Gilman, I. 2006. "From Marginalization to Accessibility: Classification of Indigenous Materials". Master's Thesis, Pacific University.

Godbold, Natalya. 2009. "User-Centered Design vs. 'Good' Database Design Principles: a Case Study, Creating Knowledge Repositories for Indigenous Australians". *Australian Academic & Research Libraries* 40, no. 2: 116-31. https://doi.org/10.1080/00048623.2009.10721390

Gosart, Ulia. 2021. "Indigenous librarianship: Theory, practices, and means of social action". *IFLA Journal* 47, no. 3: 293–304. https://doi.org/10.1177/0340035221991861

Gracioso, Luciana de Souza, Luzia Sigoli Fernandes Costa and Omar Lopes Silva. 2023. "Desafios da Organização do Conhecimento para a Representação Temática dos Saberes Originários em Sistemas de Informação: Reflexões Preliminares". In *Organização e Representação do Conhecimento em Diferentes Contextos: Desafios e Perspectivas na Era da Datificação,* edited by Natália Bolfarini Tognoli, Ana Cristina de Albuquerque e Brígida Maria Nogueira Cervantes, 615-21. Londrina: ISKO-Brasil.

Green, Rebecca. 2015. "Indigenous Peoples in the U.S., Sovereign Nations, and the DDC". *Knowledge Organization* 42, no. 4: 211-21. https://doi.org/10.5771/0943-7444-2015-4-211

Hajibayova, Lala and Wayne Buente. 2017. "Representation of Indigenous Cultures: Considering the Hawaiian Hula". *Journal of Documentation* 73, no. 6: 1137-48. https://doi.org/10.1108/JD-01-2017-0010

IBGE (Instituto Brasileiro de Geografia e Estatística). 2022. "Destaques". *Panorama* (website). https://censo2022.ibge.gov.br/panorama/?utm_source=ibge&utm_medium=home&utm_campaign=portal

Kam, D. Vanessa 2007. "Subject Headings for Aboriginals: The Power of Naming". *Art Documentation: Bulletin of the Art Libraries Society of North America* 26, no. 2:18–22. https://www.jstor.org/stable/27949465

Knight, F. Tim. 2019. *Words and Worldviews: Decolonizing Description*. Toronto: York University.

Lawson, Kimberley L. 2004. "Precious Fragments: First Nations Materials in Archives, Libraries and Museums". MA thesis, University of British Columbia.

Lee, Deborah. 2011. "Indigenous Knowledge Organization: a Study of Concepts, Terminology, Structure and (mostly) Indigenous Voices". *Partnership: The Canadian Journal of Library and Information Practice and Research* 6, no. 1: 1-33. https://doi.org/10.21083/partnership.v6i1.1427

Lee, Joel M. 1976. "E. Wyndham Hulme: a Reconsideration". In *The Variety of Librarianship: Essays in Honour of John Wallace Metcalfe*, edited by W.B. Rayward, 101-13. Sydney: LAA.

Lee, Tamara, Sarah Dupont and Julia Bullard. 2021. "Comparing the Cataloguing of Indigenous Scholarships: First Steps and Findings". *Knowledge Organization* 48, no. 4: 298-306. https://doi.org/10.5771/0943-7444-2021-4-298

Levinson, Sara. 2023. *The Latin American and Indigenous Peoples of the Americas funnel.* Paper presented at SALALM 68.

Lilley, Spencer C. 2015. "Ka Pō, Ka Ao, Ka Awatea: The Interface between Epistemology and Māori Subject Headings". *Cataloging & Classification Quarterly* 53, no. 5-6: 479-95. https://doi.org/10.1080/01639374.2015.1009671

Littletree, Sandra. 2019. "Indigenous Systems of Knowledge at the Intersections of the LIS Field and Academia". Paper presented at the International Indigenous Librarians' Forum.

Littletree, Sandra, Miranda Belarde-Lewis and Marisa Duarte. 2020. "Centering Relationality: a Conceptual Model to Advance Indigenous Knowledge Organization Practices". *Knowledge Organization* 47, no. 5: 410-26. https://doi.org/10.5771/0943-7444-2020-5-410

Littletree, Sandra and Cheryl A. Metoyer. 2015. "Knowledge Organization from an Indigenous Perspective: The Mashantucket Pequot Thesaurus of American Indian Terminology Project". *Cataloging & Classification Quarterly* 53, no. 5-6: 640-57. https://doi.org/10.1080/01639374.2015.1010113

MacDonell, Paul, Reiko Tagami and Paul Washington.2003. "Brian Deer Classification System". http://web.archive.org/web/20120520155038/http://www.slais.ubc.ca/courses/libr517/02-03-wt2/projects/deer/index.htm

Maware, Munyaradzi. 2012. *The Struggle of African Indigenous Knowledge Systems in an Age of Globalization: A*

Case for Children's Traditional Games in South-Eastern Zimbabwe. Bamenda: Langaa RPCIG.

Moorcroft, Heather. 1993. "The Construction of Silence". *The Australian Library Journal* 42, no. 1: 27-32. https://doi.org/10.1080/00049670.1993.10755624

Moorcroft, Heather. 1994. "Reflections on Constructing an Aboriginal and Torres Strait Islander Thesaurus". *Cataloguing Australia* 20, no. 4: 108-11.

Moorcroft, Heather. 1997. *Aboriginal and Torres Strait Islander Thesaurus*. Canberra: National Library of Australia.

Moraes, Miriam Gontijo de. 2023. "O desafio da organização da etnoliteratura: o caso do glossário cecolonial de Macunaíma". In *Organização e Representação do Conhecimento em Diferentes Contextos: Desafios e Perspectivas na Era da Datificação,* edited by Natália Bolfarini Tognoli, Ana Cristina de Albuquerque and Brígida Maria Nogueira Cervantes, 608-14. Londrina: ISKO-Brasil.

Moulaison-Sandy, Heather and Jenny Bossaller. 2017. "Providing Cognitively just Subject Access to Indigenous Knowledge through Knowledge Organization Systems". *Cataloging & Classification Quarterly* 55, no. 3: 129-52. https://doi.org/10.1080/01639374.2017.1281858

Nakata, Martin and Marcia Langton, eds. 2005. *Australian Indigenous Knowledge and Libraries.* Canberra: Australian Academic and Research Libraries.

Nyitray, Kristen J. and Dana Reijerkerk. 2021. "Searching for Paumanok: a Study of Library of Congress Authorities and Classifications for Indigenous Long Island, New York". *Cataloging & Classification Quarterly* 59, no. 5: 409-41. https://doi.org/10.1080/01639374.2021.1929627

Olson, Hope A. 1999. "Cultural Discourses of Classification: Indigenous Alternatives to the Tradition of Aristotle, Durkheim and Foucault". In *ASIS SIG/CR Classification research workshop: Proceedings of the 10th ASIS SIG/CR Classification Research Workshop 31 october 1999 Washington DC, USA,* edited by Hanne Albrechtsen and Jens-Erik Mai, 1107-24.

Olson, Hope A. and D. B. Ward. 1998. "Charting a Journey across Knowledge Domains: Feminism in the Decimal Dewey Classification". In *Structures and Relations in knowledge organization: Proceedings of the Fifth international ISKO Conference 25-29 August Lille, France*, edited by Widad Mustafa El Hadi, Jacques Maniez and Steven A. Pollitt. Advances in Knowledge Organization 6. Würzburg: Ergon, 238-44. Silver Spring: American Society for Information Science.

Oyarce, Ana María, Fabiana Del Popolo and Jorge Martínez Pizarro. 2009. "International Migration and Indigenous Peoples in Latin America: The Need for a Multinational Approach in Migration Policies". *Revista Latinoamericana de Población* 3, no. 4-5: 143-63. http://www.redalyc.org/articulo.oa?id=323827368007

Pettitt, Karl and Erin Elzi. 2023. "Unsettling the Library Catalog: A Case Study in Reducing the Presence of 'Indians of North America' and Similar Subject Headings". *Library Resources and Technical Services* 67, no. 2: 4-12. https://doi.org/10.5860/lrts.67n2.4

Peyser, Alexia and Juan Chackiel. 1994. "La población indígena en los censos de América Latina". In *Estudios Sociodemográficos de Pueblos Indígenas,* 27-48. Santiago de Chile: CELADE.

Silva, F. R. 2023. "Lista de cabeçalhos de assunto e tesauros: discussões e iniciativas à luz da justiça social e equidade". Master's Thesis, Universidade de São Carlos.

Simões, M. G. M. 2010. "A representação de etnia e sua evolução na Classificação Decimal Universal". PhD diss., Universidad de Salamanca.

Simpson, Sally. 2005. *Te Ara Tika: Guiding Words: Phase 3 Research Report.* [Wellington, NZ]: [Māori Subject Headings Steering Committee].

Srinivasan, R. 2002. "Village Voice: Expressing Narrative through Community-Designed Ontologies". PhD diss., Massachusetts Institute of Technology.

Srinivasan, Ramesh. 2007. "Ethnomethodological Architectures: Information Systems Driven by Cultural and Community Visions". *Journal of the American Society for Information Science & Technology* 58, no. 5: 723-33. https://doi.org/10.1002/asi.20544

Steward, Julian H. 1949. "The Native Population of South America". ln *Handbook of South American Indians*, organized by Julian H. Steward, 655-68. Washington: Government Printing Office.

Swanson, Raegan. 2015. "Adapting the Brian Deer Classification System for Aanischaaukamikw Cree Cultural Institute". *Cataloging & Classification Quarterly* 53, no. 5-6: 568-79. https://doi.org/10.1080/01639374.2015.1009669

Szekely, Chris. 1997. *Te Ara Tiki: Guiding Voices; Maori Opinion on Libraries and Information Needs.* Wellington, New Zealand: New Zealand Library and Information Association.

Taylor, Andrew, Gary Johns, Gregory Williams and Malinda Steenkamp. 2016. "The Problem of Indigenous Migration in the Globalised State". In *Demography at the Edge: Remote Human Populations in Development Nations*, edited by Dean Carson, Rasmus Ole Rasmussen, Prescott Ensign, Lee Huskey and Andrew Taylor, 163-88. New York: Routledge.

Thorpe, Kirsten and Monica Galassi. 2014. "Rediscovering Indigenous Languages: the Role and Impact of Libraries and Archives in Cultural Revitalisation". *Australian Academic and Research Libraries* 45, no. 2: 81-100. https://doi.org/10.1080/00048623.2014.910858

Tomren, Holly. 2004. *Classification, Bias and American Indian Materials* (Unpublished research paper). San Jose: San Jose State University.

Unesco (United Nations Educational, Scientific and Cultural Organization). 2001. *Universal Declaration on Cultural Diversity*. https://www.ohchr.org/en/instruments-mechanisms/instruments/universal-declaration-cultural-diversity

Velazco-Ortiz, Matilde Laura. 2023. "Indigenous Migration in Latin America". In *The Routledge History of Modern Latin American Migration*, coordinated by Andreas E. Feldmann, Xóchitl Bada, Durand Jorge y Stephanie Schütze, 186-200. New York: Routledge.

Webster, Kelly and Ann Mary Doyle. 2008. "Don't Class me in Antiquities! Giving Voice to Native American Materials". In *Radical Cataloging: Essays at the Front*, edited by K. R. Roberto, 189-97. Jefferson, NC: Greenwood.

Wise, Mary and Sarah R. Kostelecky. 2018. "Respecting the Language: Digitizing Native American Language Materials". *Digital Library Perspectives* 34, no. 3: 200-14. https://doi.org/10.1108/DLP-02-2018-0006

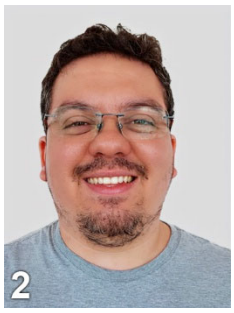
Yeh, Thomas Yen-Ran. 1971. "The Treatment of the American Indian in the Library of Congress E-F Schedule". *Library Resources and Technical Services* 15, no. 2: 122–28.

Young, Mary L. and Dara L. Doolittle. 1994. "The Halt of Stereotyping: When Does the American Indian Enter the Mainstream?" *Reference Librarian* 22, no. 47: 109-19. https://doi.org/10.1300/J120v22n47_09

Organization and Representation of Indigenous Scientific Production: A Case Study on the Institutional Repository in Brazil

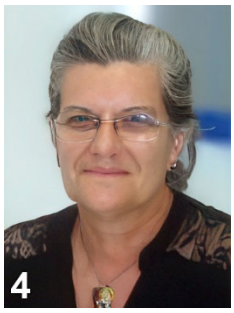
Caroline Periotto*, Felipe Arakaki **, Jair de Jesus Massa ***,
Luzia Sigoli Fernandes Costa ****, and Luciana de Souza Gracioso *****

* *** **** ***** Federal University of São Carlos (UFSCar), Brazil
** Federal University of Brasília (UnB), Brazil
* carol@ufscar.br, ** fe.arakaki@gmail.com, *** jair@estudante.ufscar.br,
**** luziasigoli@ufscar.br, ***** luciana@ufscar.br



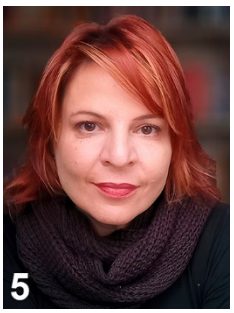
Caroline Periotto holds a Master's degree in Science, Technology, and Society from UFSCar. Librarian at the Federal University of São Carlos (UFSCar). Works with the UFSCar Institutional Repository.

Felipe Arakaki holds a PhD and a master's degree in Information Science from the São Paulo State University (UNESP) and bachelor's degree in Library Science from the São Paulo State University (UNESP). Professor of the Library Science course at the University of Brasília (UnB) and the Postgraduate Program in Information Science (PPGCI) at UFSCar.



Jair de Jesus Massa is an indigenous student of Library and Information Science at UFSCar. CNPq Scientific Initiation Scholarship recipient. *Desana* ethnicity.

Luzia Sigoli Fernandes Costa holds a PhD in Information Science from the São Paulo State University (UNESP) and a bachelor's degree in Librarianshi. Professor of the Library Science course at the UFSCar and the Postgraduate Program in Science, Technology and Society at UFSCar.



Luciana de Souza Gracioso is a postdoctoral researcher in Information Science at the University of Coimbra. She holds a Ph.D. in Information Science from the Fluminense Federal University /Brazilian Institute of Information in Science and Technology (IBICT). Associate Professor at UFSCar. CNPq Productivity Fellow/2.

Periotto, Caroline, Felipe Arakaki, Jair de Jesus Massa, Luzia Sigoli Fernandes Costa and Luciana de Souza Gracioso. 2024. "Organization and Representation of Indigenous Scientific Production: A Case Study on the Institutional Repository in Brazil". *Knowledge Organization* 51, no. 8: 642-659. 40 references. DOI:10.5771/0943-7444-2024-8-642.

Abstract: The 2019 United Nations General Assembly declared that the years of 2022 to 2032 will encompass the International Decade of Indigenous Languages. Within this context, information Science (IC) and knowledge organization (OC) play a central role in the construction of policies and actions aimed at preserving indigenous languages. As such, our research aimed to diagnose, make propositions, and run preliminary tests in an Institutional Repository (IR) of a Brazilian Federal University. More specifically, this work regards the inclusion of metadata to represent indigenous scientific production. The main operations in this study included the creation and adaptation of metadata fields in the Dublin Core scheme. These were to specifically indicate indigenous collective authorship, indigenous names in an authorship, summaries in indigenous language, keywords in indigenous language, and title in indigenous language. The methods were literature review, applied research, and an experimental research. The implementations made in RI enabled, even if on a preliminary basis, the institution's repository as a robust instrument for archiving and

accessing scientific information. It is also a means for preservation, visibility, appreciation, and respect for indigenous languages and knowledge, which are currently threatened.

Received: 29 June 2024; Revised: 2 August 2024; Accepted: 15 August 2024.

Keywords: indigenous scientific production; organization and representation of indigenous knowledge; metadata; digital repositories.

1. Introduction

Studies dedicated to Institutional Repositories (IR) have received a growing interest from the field of Information Science, particularly to enhance the process of dissemination, preservation, retrieval, and access to information. A significant amount of research in the field of Information and Knowledge Organization has been dedicated to the qualification of document representation processes in Institutional Repositories. This is a process that directly impacts the way in which registered scientific knowledge can be accessed by society.

Nevertheless, the policies that generally guide representation in repositories still deserve attention as to be implemented and provide more inclusive protocols for cataloging, indexing, and making use of language processes in the representation and description of content. The qualification of these processes in Institutional Repositories becomes essential at a time when, especially during the last decade in Brazil, there has been a growing increase in the participation of indigenous students in the country's public universities. Students who, both at undergraduate and graduate levels, begin to produce and publish scientific knowledge which is then made available by the Institutional Repositories of these universities.

An entire mode of knowledge production has recently been proposed based on indigenous worldviews towards scientific knowledge, which considers multidimensional concepts, intersectional methodologies, collective and community authorship, as well as keywords assigned to represent scientific knowledge with ancestral meanings. These are important elements in validation criteria so that indigenous scientific knowledge constructed in Brazilian public universities can be recorded, archived, recovered, and made accessible in a manner consistent with its mode of production. In this context, the research question that this article aims to answer is: How can the scarcity of library information systems that enhance the descriptive and thematic representation of Indigenous knowledge produced in Brazilian public universities be addressed through the analysis and proposal of "decolonial metadata" structures in academic Institutional Repositories?

Given this briefly exposed scenario, our research aimed to diagnose, propose, and run preliminary tests in the Institutional Repository of the Universidade Federal de São Carlos (São Carlos Federal University, hereforth noted as UFSCar). These efforts aim at the inclusion and adaptation of metadata fields for the following – indigenous collective authorship, indigenous name in an authorship, summary in indigenous language, keywords in indigenous language, and title in indigenous language. This seeks to meet some of the CARE Principles oriented towards Indigenous Data Governance, which was proposed by the Global Indigenous Data Alliance and were established during the International Data Week and Research Data Alliance Plenary (Vidotti et al. 2021).

As such, the article presents an overview of affirmative action policies for the inclusion of the indigenous population at UFSCar, thus contextualizing when the policies emerged in Brazil and when they were in fact incorporated into the University. A bibliographical survey was subsequently carried out on indigenous representation and Knowledge Organization Systems, which provided support for understanding the importance of this theme in this particular field. Some of the laws guaranteeing the use of indigenous languages in Brazil are subsequently presented, thus reinforcing the importance of document representation processes, especially in Digital/Institutional Repositories. Digital Repositories are contextualized as a tool that guarantees the preservation, recovery, and access to scientific knowledge. The emergence of the UFSCar Institutional Repository is then contextualized, and finally, the presentation of the actions and results of the Dublin Core field and metadata inclusion for representing the indigenous language.

2. Affirmative action policies for the inclusion of the indigenous population at University

According to the Institute of Geography and Statistics (IBGE) 2010 Brazilian census, the indigenous population is almost 900 thousand people, with the country's total population being approximately 212.7 million people. There are 305 ethnicities who speak more than 270 different languages in the Brazilian territory. When it comes to enabling access to educational resources for the indigenous peoples of this country; however, there is still a lot to be done despite some important achievements.

Brazil held 67 federal universities and 39 state universities offering an average of 390 thousand openings in public

higher education in 2020. In 2012, Law No. 12,711 was enacted (Brazil 2012). It established a reservation of places in higher education federal institutions and technical courses for students coming from public schools. This Law enabled low-income, black, mixed-race, indigenous, and people with disabilities to have better conditions to access education. Its Article 3 states that "places must be allocated for people self-declared as black, mixed race, indigenous and quilombola, as well as people with disabilities in each federal higher education institution". The offer for indigenous entrance exams began in 2001 in Brazil, and in 2006 at UFSCar.

In addition to the quotas determined by law, 32 institutions offer entrance exams aimed exclusively at indigenous candidates – UFSCar among them. Data collected by the Higher Education Census of the Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) indicate that 57,706 indigenous people were enrolled in higher education in 2018. This is a 695% increase when compared to the initial implementation years of affirmative action policies and indigenous entrance exams in Brazil, which began in 2010 (Luciano and Amaral 2021).

UFSCar is a Brazilian federal university founded in 1968 with 4 campuses spread across the interior of the state of São Paulo offering a total of 64 undergraduate courses and 52 postgraduate programs. It was only in 2008 that there was an official record of the presence of indigenous peoples at this university. This fact itself attests to the results of implementing the Indigenous Entrance Exam by the university. Since the first edition of this entrance exam, more than 60 different indigenous ethnicities have passed through the university. The year of 2023 saw the largest number of registrations for this entrance exam, which is currently carried out in partnership with Universidade de Campinas (Unicamp). There were 3,480 applications for the 130 openings at UFSCar and the 130 places at Unicamp. There are approximately 400 indigenous students with active enrollment at UFSCar at the moment. In 2016, the affirmative action policy aimed at postgraduate studies was approved at this institution. Due to the recent nature of this implementation, the presence of indigenous students in graduate studies is small, but growing. It is safe to affirm that scientific production produced by indigenous peoples will only increase when considering the implemented policies and the systematic increase in the entry of indigenous students into Brazilian undergraduate and postgraduate studies. This creates a demand that the storage systems and organization of academic scientific production be readily represented and recovered, thus giving greater visibility to indigenous languages, ethnicities, and authorship.

It is thus agreed that public universities – especially in countries affected by colonizing processes – cannot be another place for the Coloniality of Being (Nelson Maldonado-Torres and Walter Mignolo), Power (Aníbal Quijano), or Knowledge (Santiago Castro-Gômez) and must then be constituted as Pluriversity of Knowledge (Boaventura de Souza Santos). The justifications for actions aimed at expanding and qualifying diverse participation, especially that of the indigenous population in scientific production spaces, are multiple. These include issues of an environmental nature and the preservation of ways of life, social relations, society, and nature – which according to the conception of indigenous peoples, cannot be seen separately.

As such, libraries, archives, museums, and different information units occupy a central place in promoting cognitive justice in countries victims of colonial oppression. For example, the Indigenous Matters Section was created within the International Federation Library Association (IFLA) with the aim of supporting culturally responsive and effective services to indigenous communities around the world. The section promotes international cooperation in libraries, culture, knowledge, and information services to meet the cultural, linguistic, and community needs of indigenous communities. Furthermore, it encourages indigenous leadership, exchange of experiences, education, training, and research.

3. Indigenous languages and knowledge organization systems

Significant academic work has also been produced, especially within the scope of Knowledge Organization in an attempt to mitigate the side effects from the imposition of systems, languages, models of representation, and information retrieval which can either optimize or render indigenous knowledge invisible.

An important work to mention, as a starting point, is Doyle's (2013) doctoral thesis titled Naming, claiming, and (re)creating: Indigenous knowledge organization at the cultural interface. The thesis will demonstrate that prevailing knowledge organization systems (KOS) within libraries frequently encounter substantial difficulties in the representation and organization of documents containing Indigenous content. These difficulties underscore a deficiency in traditional approaches, which may inadequately address the complexities and nuances inherent in Indigenous perspectives. Nonetheless, conceptual, theoretical, and methodological approaches derived from Indigenous knowledge have the potential to offer innovative directions for the design and development of KOS.

Recent research developed by Littletree et al. (2020) addresses the need to create a Conceptual Model to Advance Indigenous Knowledge Organization Practices. The authors expose the limitations of traditional Euro-American approaches to Knowledge Organization (KO) when dealing with indigenous issues. They indicated an urgency for professionals working with KO to understand indigenous per-

spectives on an epistemological level. Within this context, another work was published in the Proceedings from the North American Symposium on Knowledge Organization by Lee et al. (2021) entitled 'Comparing the Cataloguing of Indigenous Scholarships: First Steps and Findings'.In this work, an analysis of terms within the Library of Congress subject heading was made on the term "Indian". The research highlights two main issues related to the continued use of current LCSH terms. First, they are ambiguous and limit the effectiveness of an institutional catalog. Secondly, these terms do not reflect how Indigenous Peoples, Nations, and Communities in North America.

Villanueva (2021), in 'Language Subject Access to Indigenous Materials: The Philippine Cordillera Case', examines the limitations of the Library of Congress Subject Headings (LCSH) in assigning access terms to Indigenous materials from the Cordillera region of the Philippines. The article identifies a significant challenge: many Indigenous terms lack direct equivalents in English or other Western languages, which impedes both the cataloguing and efficient retrieval of these materials. Additionally, issues such as variations in spelling and local nomenclature further complicate the indexing process. To address these challenges, Villanueva proposes several solutions and steps aimed at improving the searchability and accessibility of these materials, striving to overcome existing barriers and enhance information access. prefer to represent themselves as individuals or collectives. Andersen and Skouving (2017) report on the challenges of incorporating the specificities of the Maõri language into Organization of Knowledge systems in the book 'Classification and decolonisation of Maõri Subjects', published in The Organization of Knowledge: Caught Between Global Structures and Local Meaning.

Another work that problematizes bibliographic classification systems in terms of indigenous issues was developed by Green (2015), entitled 'Indigenous Peoples in the U.S., Sovereign Nations, and the DDC'. The author investigated the DDC in editions 16 to 23, and demonstrated some actions already developed on code tables in order to minimize inconsistencies in the classification structures on indigenous peoples. The research developed by Adler (2016), entitled 'The Case for Taxonomic Reparations', recognizes that static structures in bibliographic classification systems end up highlighting processes through which violence became systemic, especially in relation to eugenics. As a result, it proposes the creation of reparative taxonomies and uses the cases #BlackLivesMatter and the Transgender Digital Archive to exemplify possibilities in the construction and use of liberatory descriptive standards that are transparent, hold creators accountable, and invite user participation in the co-creation of records.

There have also been important decolonization movements in relation to collections within museums and archives. Turner (2016), in her work on Organizing Knowledge in Museums, calls attention to the practical and intellectual issues raised when other forms of knowledge meet museums systems (472). Furthermore, the author contextualizes the study of museums within Foucaultian thought and other concepts that help understand the standardization of systems via institutional repositories, as these are increasingly being digitized and disseminated across varied communities. In his article Ethical Issues from Decolonial Practices in Knowledge Organization: The Case of Indigenous Collections in Världskulturmuseet, Sundström (2023) presented the case of the Museum of World Culture (Varldskulturmuseet), which has revised its knowledge organization system – especially with regard to the classification and use of terminologies on indigenous people in their database known as Carlotta. The research confirmed that the museum used inappropriate and obsolete terms to describe some of the objects in the indigenous collection and concluded that indigenous self-determination in cultural institutions is essential so that ethical issues related to the creation of a knowledge organization system in a museum can be respected. In the archival field, the report by Gilliland (2012) 'Contemplating Co-creator Rights in Archival Description' debated the ethics and rights of co-creators while seeking to highlight how indigenous conventions can be incorporated into archival descriptions as resources to recognize the co-creation of documents. The author concluded that it is necessary to rethink archival descriptive practices and standards so that they can consider ethical and power differences that may manifest within the archival multiverse.

Given this briefly announced context, this research addresses the research gap on library information systems for recording and retrieving information that can enhance the descriptive and thematic representation of indigenous knowledge produced in Brazilian public universities while considering their original cosmogony as structuring elements of metadata representation. Among these systems, the research is dedicated to analyzing and making propositions for structuring "decolonial metadata" in Institutional Repositories.

Studies dedicated to Institutional Repositories have aroused growing interest in the field of Information Science, and it is possible to affirm that it composes the majority scientific knowledge produced on the subject in the world (Bazilio 2022). Specifically within this field, a significant production of research has been developed in Knowledge Organization Literature Classification Systems (CSKOL) (Alves et al. 2022; Fujita 2022; Fujita et al. 2022; Fujita et al. 2023; Fujita and Panuto 2024; Fujita et al. Panuto 2024). These are dedicated to the qualification of indexing processes of archived content in Institutional Repositories, since this process directly impacts the way in which

registered scientific knowledge can be accessed by society. Specifically in relation to UFSCar's IR, Freitas (2019) analyzed the process of representing issues made by authors of scientific works when archiving them. The research conclusively indicated that policies, and even metadata structures, still deserve more attention in order to provide opportunities for more inclusive protocols in cataloging, indexing and use of language in representation processes and content description.

Preliminary studies on indigenous representation at UFSCar have identified and discussed the gap that gave rise to this research proposition. It was presented during ISKO Brasil 2023 in Londrina – Paraná – and later published in Anais do Congresso (Gracioso et al. 2023) and presented at the VII Information, Data, and Technology Workshop, held in 2024 in Porto Velho – Rondônia (Periotto et al. 2024).

4. Metadata for the preservation and appreciation of knowledge from original peoples – legal starting points

Brazil has only one official language – Portuguese. The processes and paths that shaped this institutionalization are complex and result from radical colonization strategies by the Portuguese nation over the population of Pindorama. Records indicate that there were more than 1000 languages being spoken in Brazilian territory before the colonization process. From 1500 to 1750, many of the surviving languages still coexisted and survived alongside the imposed Portuguese language, with the Tupi language predominating in the coastal regions and Nheengatu in the Amazon. A Portuguese Royal Provision prohibited the use of Tupi in 1757 and established Portuguese as the official country's language. In 1759, the use of original languages by indigenous peoples was criminalized. It was only 200 years later that social organization, customs, languages, beliefs, and indigenous traditions were recognized with the publication of the 1988 Constitution of the Federative Republic of Brazil (article 231).

Currently, there has been a growing investment in the country for the socio-political inclusion of indigenous peoples in different instances. This includes discussion about the right to use original languages, as well as the set of legal standards that regulate the use of languages and the exercise of linguistic rights in Brazil. These can be accessed via the Brazilian Repository of Linguistic Legislations (RBLL), developed in partnership with the Institute for Research and Development in Linguistic Policy (IPOL). More than 35 years after the publication of the 1988 Constitution, the Brazilian Federal Supreme Court launched the first Brazilian Constitution translated into the indigenous Nheengatu language *"MUNDU AS TURUSU" WAÁ ÜBÊUWA MAYÉ MÍRA ITÁ UIKÚ ARÃMA PURÃGA IKÉ BRAZIU UPÉ*".

Of the more than 270 indigenous languages in Brazil, only 13 are currently recognized as co-official in 10 Brazilian municipalities (of the more than 5,000 municipalities in the country). The co-official indigenous languages and the respective laws that regulate their use are described in the Table 1, prepared by the Institute for Research and Development in Linguistic Policy.

In addition to the right to use their languages, advances have also been made in the rights of indigenous peoples to self-determine. These were developed both internationally via the 169th International Labor Organization Convention (ILO) of 1989, and nationally via the 1988 Federal Constitution. More recently, this has also been achieved via Resolution No. 3 of 2012 of the National Council of Justice, which guarantees original communities the right to rectify Brazilian names for indigenous people in the Civil Birth Registry (RCN) as a record made in the Civil Registry Offices of Natural Persons – provided for and regulated by Law 6,015/73. Article 2 of the joint Resolution CNJ/CNMP nº 03/2012 ensures that in Civil Birth Registries "the indigenous name of the registrant, of his/her free choice, must be entered at the request of the applicant (...)". Likewise, indigenous people who want to correct their already registered names, or change them to add the people or ethnicity, can request such a change from the registry offices. Furthermore, according to Joint Resolution No. 3, of April 2012, an indigenous ethnicity can be entered as a surname, if a person so chooses. The village of origin of the indigenous person and that of their parents may be included as information regarding their respective places of birth, along with the municipality of birth. According to Art. 42, VI of TSE Resolution No. 23,659/2021 and the ELO System (Brazilian voter registration), it is currently possible for the person to identify as indigenous, enabling one to declare their identity, ethnicity, and indigenous language (self-declaratory information) in the system.

As such, it is possible to see that actions have been taken – largely due to the demands from indigenous peoples for their rights – in an attempt to minimize the side effects of colonization processes that almost extinguished indigenous bodies, places, languages, and cultures in Brazil. This ensures that we will not only be doing due cognitive and authorial justice, but also practicing the Law when creating metadata for adequate representation of indigenous scientific production in open and free information systems such as Institutional Repositories.

Language	Municipality	Law – Year
Baniwa	São Gabriel da Cachoeira (AM)	Law n. 145/2002
Neengatu	São Gabriel da Cachoeira (AM)	Law n. 145/2002
Tukano	São Gabriel da Cachoeira (AM)	Law n. 145/2002
Ianomami	São Gabriel da Cachoeira (AM)	Law n. 0084/2017
Guarani	Tacuru (MS)	Law n. 848/2010
Akwê-Xerente	Tocantínea (TO)	Law n. 411/2012
Macuxi	Bonfim (RR)	Law n. 211/2014
	Cantá (RR)	Law n. 281/2015
Wapichana	Bonfim (RR)	Law n. 211/2014
	Cantá (RR)	Law n. 281/2015
Mebêngôkre/Kayapó	São Felix do Xingu (PA)	Law n. 571/2019
Tenetehara/Guajajara	Barra do Corda (MA)	Law n. 900/2020
Tikuna	Santo Antônio do Içá (AM)	Law n. 298/2020
Tupi-nheengatu	Monsenhor Tabosa (CE)	Law n. 13/2021
Terena	Miranda (MS)	Law n. 1.382/2017 Law n. 1.417/2019

Table 1. Co-official indigenous languages in Brazilian municipalities.
Source: List of co-official languages in Brazilian municipalities | IPOL.

5. Digital Repositories: preservation, recovery, and access to scientific knowledge

The open access movement has brought an expansion and popularization of digital academic and technical-scientific work repositories, among other types of materials available. This operates mainly in universities and research centers. "The expression 'digital repositories' [...] is used to describe the various types of data provider applications that are intended for the management of scientific information, constituting alternative routes of scientific communication." (Leite 2009, 19). According to Torino (2017, 94):

Digital repositories (RDs) are open and interoperable information systems intended for the management of scientific and academic information. These are capable of storing files of different formats, and constitute alternative ways of scientific communication while increasing production visibility.

The *Conselho Nacional de Arquivos* (2023, 12, translated by the authors) reiterates that the digital repository is "a system that supports the management of digital materials for as long as necessary, and is made up of hardware, software, and metadata, as well as an organizational infrastructure with normative and technical procedures" (p.9). Gonçalez (2017, 79-82) adds that digital repositories are made up of digital collections that can be built in different ways and purposes, but that the action of designating the creation of a digital repository favors the storage of a large number of documents. These must be capable of managing and storing collections of digital objects for long periods with appropriate access conditions.

Institutional repositories are digital platforms maintained by academic institutions such as universities, research institutes, libraries, and documentation centers. These repositories are designed to store, preserve, and make the intellectual and scientific production by the academic community available. They generally provide a variety of content such as journal articles, theses, dissertations, research reports, research datasets, and other forms of academic production.

Institutional repositories play a crucial role in promoting open science because (i) they facilitate free and immediate access to scientific knowledge, allowing researchers, students and the general public to consult, read, and use research results without financial or technical restrictions ; (ii) provide a safe and reliable place to store and preserve research results over time, thus ensuring that scientific knowledge is not lost and remains available for future generations; (iii) visibility and impact increases by making an institution's research works available, as results can be easily found and accessed by other researchers, collaborators, and other interested parties around the world; (iv) allow institutions and researchers to comply with funding agency policies that require publicly results to be made available in open access;

(v) facilitate collaboration and networking between researchers, allowing them to share their work, ideas, and resources, promoting a more collaborative and productive research environment.

Institutional Repositories are defined based on their characteristics, application and intended objectives (Torino 2017, 95) and can be divided into institutional, thematic, archival, data, and other categories. According to Leite (2009, 21), an institutional repository is "a scientific information service in a digital and interoperable environment dedicated to managing an institution's intellectual production." Thematic repositories correspond to the management of material grouped by a common subject.

Many solutions were created in order to enable the management of digital documents for digital repositories, DSpace stands out among them. DSpace (n.d.) is a free, open source software developed by the Massachusetts Institute of Technology (MIT) library in conjunction with Hewlett-Packard (HP) in the early 2000s, but it is now maintained by LYRASIS.

Among DSpace's main features are: a web-based interface; the possibility of exporting different file formats such as text, images, and digital videos; document curation and management as each file inserted into DSpace has technical information (metadata) that allows digital management and preservation in addition to navigation and indexed items search. "Once an item is located, native Web formatted files can be viewed in a Web browser, while other formats can be downloaded and opened by a suitable application." DSpace makes it possible to build communities and collections in order to organize the collection. "A community is the highest level of the DSpace content hierarchy."

Other features of DSpace are:

- Free open source software;
- Fully customizable as to meet user needs;
- Manages and preserves all digital content formats (PDF, Word, JPEG, MPEG, TIFF files);
- Apache SOLR-based search for metadata and full-text content;
- UTF-8 support;
- Interface available in 22 languages;
- Group-based granular access control, allowing one to set permissions to individual files;
- Optimized for Google Scholar indexing.
- New functionalities are developed with each version as to improve data curation management, and the current version of DSpace is 7.6.1.

6. Methodological Procedures

The methods used to develop this research were a literature review and applied research. The latter sought to solve an immediate problem on the need to expand the metadata of the Institutional Repository with the aim of representing more inclusive information.

This research is still exploratory with elements of descriptive research, as it characterizes UFSCar's institutional environment while describing the indigenous population of this university. It has additional elements from experience reports at the starting point of the research and partial results. We relied on Daltro and Fria (2019) as it explains that the Experience Report is "A scientific narrative in post-modernity" in order to anchor the choice of this methodological resource. The authors indicate that "The documentary construction of an Experience Report implies, *a priori*, that its author(s), or at least one of them, is a participant in the real-life context under study." (234). As such, this present research is developed by agents composed of librarians and indigenous students who are directly developing, testing, validating, and applying the metadata resources proposed to the RI UFSCar. Another characterizing element of the research as an Experience Report is the fact that, as indicated in Daltro and Faria (2019, 234) "those already affected by the discourse of science can locate the power of theorization for advancement in a certain field of knowledge."

The study reports some elements of experimental research, since variables in the study's framework (Institutional Repository) were manipulated. This, in turn, allowed the testing of hypotheses raised in the study.

Even on a preliminary test, the implementations in the Institutional Repository from this research reinforced that these are configured not only as an instrument for archiving and accessing scientific information, but also as a technology for preservation, visibility, appreciation, respect, and strengthening of indigenous knowledge and languages that have historically been silenced and are now threatened.

It is worth mentioning that, in order to carry out these tests, we used the scientific production of one indigenous student linked to the Special Education Course at the São Carlos Campus. The chosen scientific work was an undergraduate research entitled "A'uwe Xavante reports and narratives on Special Education". The title in the indigenous language is "*Rówatsu 'u duré iwatsu 'u A'uwe Xavante ró-bdzanhamri rómnhõré îpe tsiré*".

7. The Institutional Repository at Universidade Federal de São Carlos

The UFSCar Institutional Repository (RI UFSCar) was approved through ConsUni Resolution No. 835, of March 4th, 2016, and officially implemented in June 2nd, 2016 (Universidade Federal de São Carlos 2016)

The Theses and Dissertations Collection was the first to populate the repository, and originated from the digital collection of the Digital Library of Theses and Dissertations at

UFSCar. The UFSCar Digital Library of Theses and Dissertations began in 2004 and used the Theses and Dissertations Electronic Publication System (TEDE), developed and maintained by the Brazilian Institute of Information in Science and Technology (IBICT). This institution aims to provide for the implementation of digital libraries of theses and dissertations in educational and research institutions and is thus integrated into the Brazilian Digital Library of Theses and Dissertations.

The RI UFSCar was developed on the free software DSpace, using Dublin Core as a metadata standard to provide online open access and visibility to digital items produced in text, image, sound, audiovisual, and other possible formats of intellectual production along with their metadata (Universidade Federal de São Carlos 2024).

As of October 20, 2017, RI UFSCar began to receive theses and dissertations via self-deposit carried out by the student who authored the work.

It was only in February 22nd of 2024 that the CO/SIBI Resolution No. 3 began to officially deal with Repository Policy

The UFSCar Institutional Repository (RI UFSCar) aims to organize, store, disseminate, and preserve the intellectual production of teaching and technical-administrative employees; undergraduate and graduate students; other researchers, and editorial labels linked to UFSCar, as well as those prepared through agreements or collaboration between UFSCar and other bodies published in authorship or co-authorship (Universidade Federal de São Carlos 2024).

The UFSCar Institutional Repository Policy also determined the self-deposit modality. It is worth mentioning that the UFSCar Institutional Repository is linked to the UFSCar Integrated Library System (SIBi-UFSCar) and managed by the Department of Scientific Production (DePC). The Scientific Production Department has a team of four librarians and technical support from the Information Informatics Secretariat (SIn) to keep the DSpace software active.

The UFSCar Integrated Library System (SIBi-UFSCar) was established in 2014 by Resolution/CoAd nº 069/2014 of November 28th, 2014 and is made up of the libraries in the four UFSCar campuses: Araras Campus Library (B-Ar); Lagoa do Sino Campus Library (B-LS); Sorocaba Campus Library (B-So); and Community Library (BCo) (Sistema Integrado de Bibliotecas da UFSCar n.d.). Its main actions are focused on administrative and information management policies aiming to support UFSCar's activities. Furthermore, one of the competencies of the UFSCar Integrated Library System (SIBi-UFSCar) is to propose policies and flows to manage the UFSCar Institutional Repository and Periodical Portal.

The first collection implemented in the UFSCar Institutional Repository – Collection of Theses and Dissertations – currently holds 10,727 dissertation and 5,145 theses. The Research Data Collection was launched in November 2019, and the deposit of research data must be carried out by one of the authors of the dataset with an active link to UFSCar. Currently the Research Data Collection has 81 deposits at RI UFSCar. The Course-Completion Research Collection (TCC) was launched in July 2020; however, in this collection the deposit of the undergraduate TCC must be carried out by the student's teacher-advisor. Currently, the Course Completion Work Collection (TCC) holds 2,759 undergraduate research works in the repository. The last collection to be implemented was the Articles Collection, launched in August 2022. The self-deposit can be carried out by teachers, researchers, administrative technicians and students with an active connection to the University. Currently, the Articles Collection has 93 deposits in the UFSCar Institutional Repository.

Future prospects for implementation at the UFSCar Institutional Repository include the Book and Specialization Course Completion Research Collections.

8. Results: Actions and results in the inclusion of the Dublin Core field and metadata for indigenous language representation

A series of issues were analyzed for indigenous languages to be represented in the scientific production at UFSCar. These are for now available at the RI UFSCar test environment, so that such actions can be eventually implemented in the repository.

The main questions, as well as the proposed actions, are explained below:

1. Creation of the field in the Dublin Core metadata scheme that will allow the insertion of metadata on the ethnicity of an indigenous authors (indigenous students and researchers) in the repository submission forms. A consultation was carried out in several institutional repositories in Brazil to check whether any metadata to describe ethnicity had previously been created; however, no example was found. For this purpose, the field "dc.contributor.authorethnicity" was created. At this point, the Dublin Core “dc.contributor” metadata was qualified, as it refers to the person who participates in the resource’s artistic or intellectual production. The flexibility of the Dublin Core enables one to qualify existing metadata, as well as to create completely new metadata. As ethnicity is associated with the person, it was considered appropriate to add an “ethnicity” field next to “dc.contributor.author”. This is an optional element as it is sensitive information, and its scope of use was established as: Field created to represent the ethnicity of the indigenous author, as shown in Figures 1 and 2.

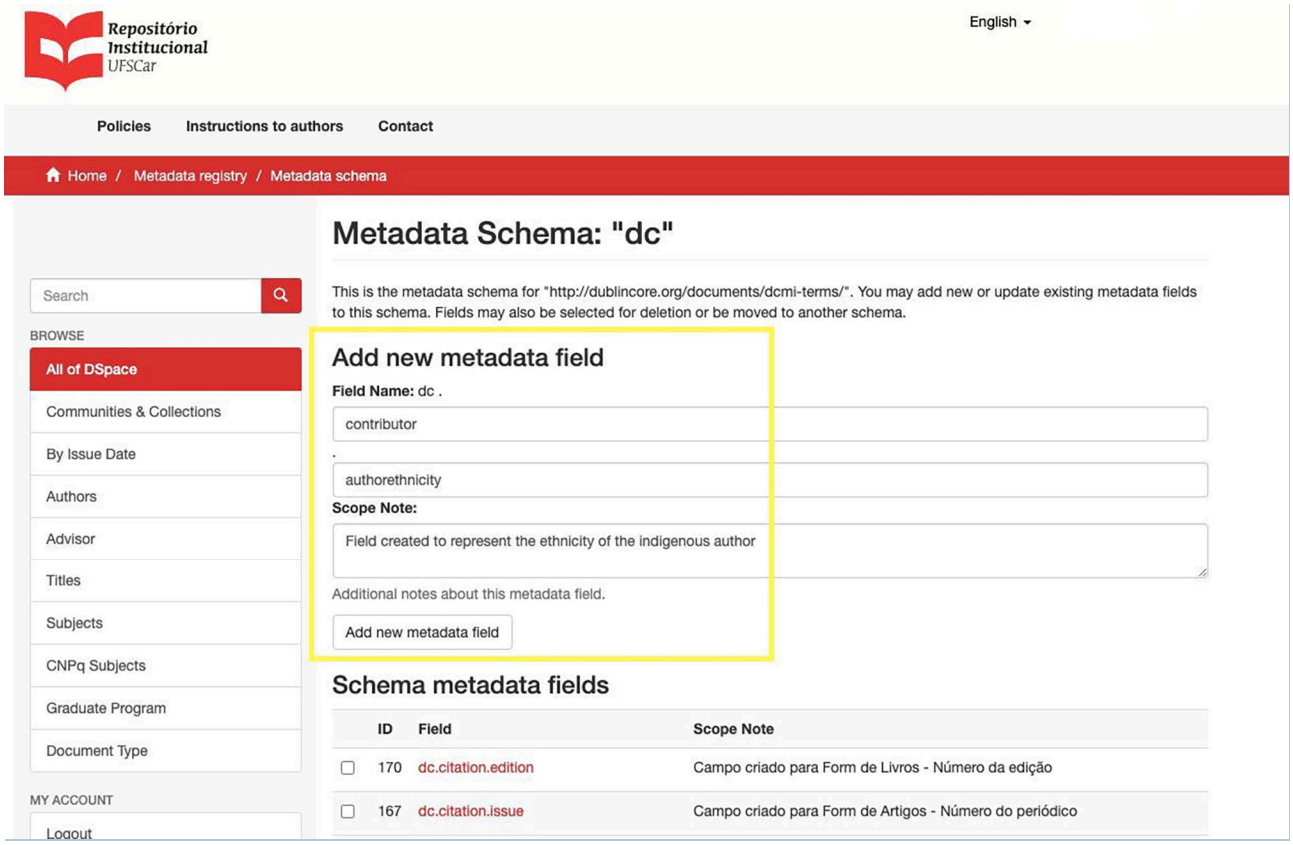


Figure 1. Inclusion of the "dc.contributor.authorethnicity" field in the Dublin Core scheme
Source: UFSCar Institutional Repository – internal pre-test interface.

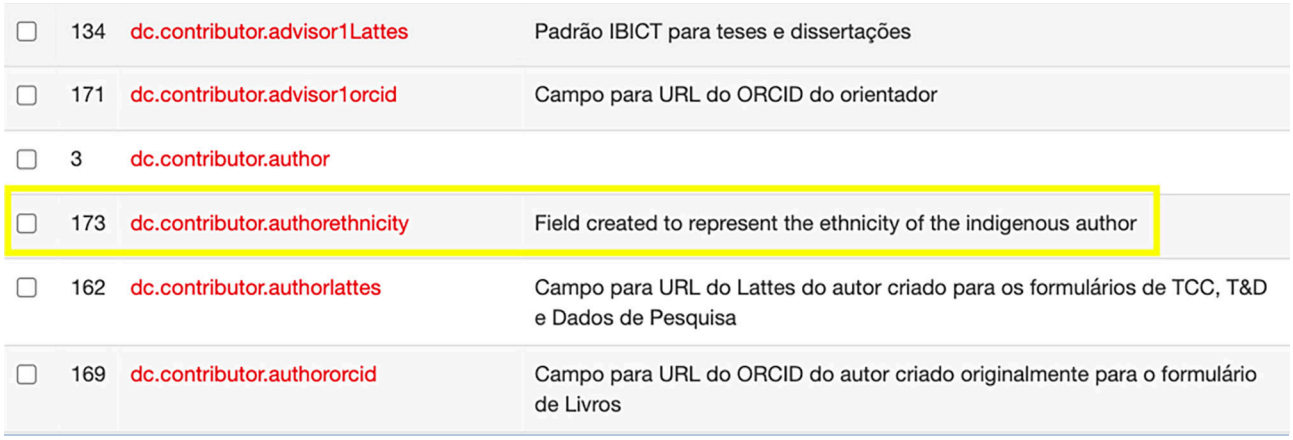


Figure 2. "dc.contributor.authorethnicity" field included in the Dublin Core scheme.
Source: UFSCar Institutional Repository – internal pre-test interface.

2. Definition of the name and description of the "dc.contributor.authorethnicity" field for item submission forms to the RI UFSCar. The field is available right after the Author field in the item submission forms, and the options for the field name and description have already been presented for consideration, as observed: (i) Ethnicity of the indigenous author or Ethnicity of the author; (ii) Enter the ethnicity of the indigenous author (without abbreviations). Use capital letters only at the beginning of words. Example: Xavante. For each ethnicity, click the Add button or Enter the author's indigenous ethnicity (without abbreviations) only in cases of admission to UFSCar through the Indige-

nous Entrance Exam. Use capital letters only at the beginning of words. Example: Guarani. Click the Add button for each ethnicity entered.

3. Identification of ethnicities already registered at UFSCar. An effort was carried out as to identify the ethnicities present at UFSCar, which resulted in a large number found. This diversity allowed the choice of metadata insertion method, where the depositor manually inserts the metadata on the ethnicity to which the author belongs for later validation by the RI UFSCar team. Figure 3 exemplifies the field and metadata referring to ethnicity.

4. Update of self-deposit manuals for all scientific production collections at RI UFSCar, with the inclusion of this new field and its guidelines. It is emphasized that this is not a mandatory field, and that it should only be filled in by indigenous students. In addition to the guidelines provided in the item submission forms, all self-deposit manuals available for consultation, as well as Instructions to authors guidance will be updated with guidelines for other fields (abstract, keyword, title in another language). Figure 4 shows the possibilities for updating one of the self-deposit manuals. It is possible to observe the respective fields in Table 2.

The "dc.contributor.author" field already provides the insertion of more than one author in all RI UFSCar collections, thus meeting demands related to indigenous collective authorship, if necessary.

It is worth highlighting that it was necessary to analyze how the language identifiers for such metadata would be inserted in addition to the field on the author's ethnicity and the guidelines for the other fields.

It is necessary to insert the language identifier for the metadata in the following fields during the validation process of items deposited at the UFSCar RI: dc.description.abstract (summary in another language); dc.language.iso (text language); dc.subject (keyword); dc.title (title); and dc.title.alternative (alternative title).

dc.type	TCC	por
dc.contributor.advisor1	Lacerda, Cristina Broglia Feitosa de	
dc.contributor.advisor1Lattes	http://lattes.cnpq.br/9468232016416725	por
dc.contributor.advisor-co1	Corsi, Adriana Maria	
dc.contributor.advisor-co1Lattes	http://lattes.cnpq.br/5553922167803121	por
dc.description.resumo	Este trabalho tem início com a minha história de vida na Aldeia Sangradouro, MT, do povo A'uwe Xavante, até chegar na Universidade Federal de São Carlos, como aluno do Curso de Licenciatura em Educação Especial. Com base em artigos sobre a história e a educação do meu povo, publicados por parentes, escrevo sobre o nosso a nossa história, nossa aldeia, a Escola São José de Sangradouro e sobre a Educação Especial. Com o desenvolvimento deste trabalho procuro conhecer como vivem as pessoas com deficiência na aldeia Sangradouro e refletir sobre a possibilidade de ajudar na integração social dessas pessoas. Por meio de conversas informais com parentes e pessoas com deficiência, apresento um pouco de suas histórias e de como vivem na aldeia. Tendo também uma filha com diagnóstico de Deficiência Intelectual, descrevo em profundidade suas relações familiares, suas atividades e o seu desenvolvimento. A partir de um plano desenvolvido para minha filha, descrevo a realização de atividades envolvendo histórias, diálogos, rotinas e convivência com a família e na aldeia.	por
dc.publisher.initials	UFSCar	por
dc.subject.cnpq	CIENCIAS HUMANAS::EDUCACAO::TOPICOS ESPECIFICOS DE EDUCACAO	por
dc.publisher.address	Câmpus São Carlos	por
dc.publisher.course	Educação Especial - EEspL	por
dc.contributor.authorethnicity	Xavante	sai

Figure 3. Inclusion of the metadata of the indigenous author's ethnicity.
Source: UFSCar Institutional Repository – internal pre-test interface.

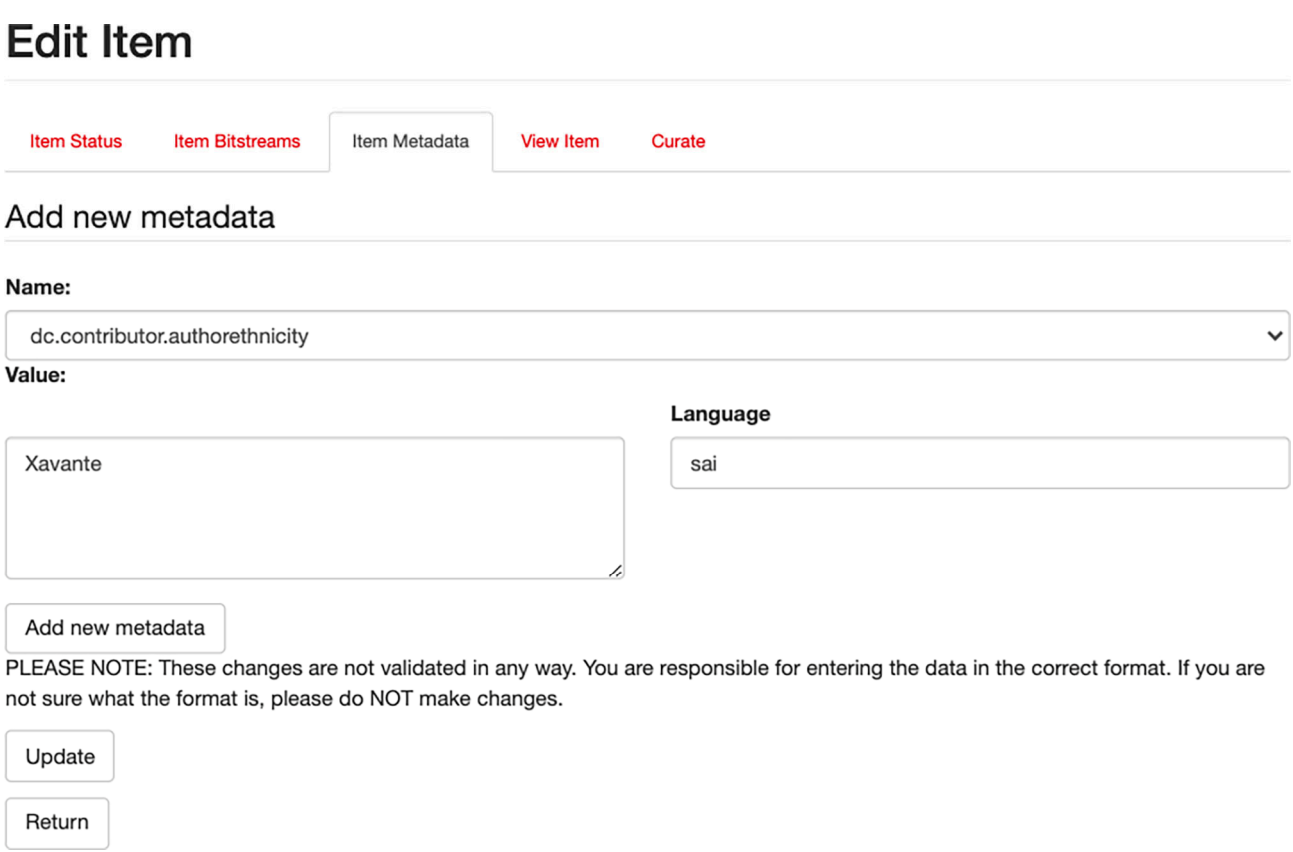


Figure 4. Inclusion of the metadata of the indigenous author's ethnicity.
Source: UFSCar Institutional Repository – internal pre-test interface.

Dublin Core field	Field name	Field description
dc.contributor.authorethnicity (field inclusion)	**Indigenous author's ethnicity**	Enter the ethnicity of the indigenous author (no abbreviations). Use capital letters only at the beginning of words. Ex.: Xukurú de Ororubá. For each ethnicity, click the Add button.
dc.contributor.author (adequacy of the field filling description)	**Author**	Enter the author's full name (no abbreviations). Use capital letters only in the initials of first and last names. For each name entered click the Add button. **In the case of a social name or indigenous name, use the name as it appears in the UFSCar systems.**
dc.description.abstract (adequacy of the field filling description)	**Abstract in a foreign language**	Provide the summary of the work in a foreign language or in an indigenous language (if applicable).
dc.subject (adequacy of the field filling description)	**Keywords**	Enter the keywords in Portuguese, in a foreign language, and in an indigenous language (if applicable). Use capital letters only at the beginning of words and in acronyms.
dc.title.alternative (adequacy of the field filling description)	**Title in another language**	Enter the title and subtitle in another language. If the title in Portuguese was entered in the previous field, enter the title in English, Spanish, indigenous language, etc.

Table 2. Creation and description of indigenous language fields in the scientific production by indigenous students.
Source: Prepared by the authors.

The language identifier at RI UFSCar follows the three-digit codes established by the ISO 639-2:1998 standard. Therefore, a consultation was carried out with the data from the Library of Congress (2024), as shown in Table 3: Codes for the Representation of Languages Names and ROSSIO Vocabularies. This follows the determinations in ISO 639-2:1998 for the best option to represent the metadata inserted in the indigenous language.

As shown in Table 3, there was an attempt to gather all language identifiers of the indigenous languages present at UFSCar; however some such as Kambeba, Lanawá, Manchineri and Wassu Cocal were not identified in the sources. As such, a language identifier of South American Indian languages was established due to the countless number of languages identified. This was made in an effort to anticipate a possible difficulty in identifying these languages when validating the work of indigenous students, and is represented by the acronym sai, which will be used for all the words and texts inserted in the RI UFSCar in any indigenous language.

5. Manual and retrospective insertion of the Dublin Core field and the mentioned metadata for items deposited in the RI UFSCar prior to the implementation of the "dc.contributor.authorethnicity" field. Mapping will be carried out to identify Course Conclusion research work, Dissertations, and Theses of indigenous authorship, so that it is possible to contact the author and manually insert the metadata for indigenous ethnicity; summary, keywords and title in the indigenous language, as shown in Figures 5, 6, and 7.

The inclusion of ethnicity metadata is a demand from indigenous students at UFSCar and will enable not only its more effective recovery, but the studies developed by them, with a more reliable description of their scientific and academic production at the university.

Figure 7 illustrates how the summary in the indigenous language stands out on the main page of items deposited at RI UFSCar.

The inclusion of all metadata presented in a course completion research contributes to the reliable representation of the indigenous ethnicity in scientific production.

9. Final Considerations

This research aimed to diagnose, make propositions, and pre-test the inclusion of metadata for the representation of indigenous scientific production in an Institutional Repository (IR) of a Brazilian Federal University. The research fulfilled its objective, demonstrating a mirroring of the system on a pre-test where metadata was implemented. It is worth mentioning that the inquiry which led to this study was guided by indigenous students at UFSCar.

The UFSCar Institutional Repository initially carried out several studies to adapt the representation established from Dublin Core in the DSpace Software. There were discussions on how the UFSCar Institutional Repository could expand its possibilities for searching and retrieving information while respecting the various institutional guidelines, policies, and legislation on the subject.

The inclusion of ethnicity metadata to facilitate the search and retrieval of information on certain groups and studies by indigenous researchers was suggested, as well as allowing the representation of ethnicity by indigenous authors in UFSCar's scientific production. Other solutions were proposed, such as enabling the inclusion of the indigenous name in the repository. The inclusion of a summary, keywords, and title in the native language was also provided with their appropriate linguistic coding, thus maintaining the integrity and consistency of the information in the repository. The use of international coding in addition to maintaining standardization guarantees efficient interoperability with other information systems.

It is important to highlight that the implementation of the ethnicity of indigenous author field, as well as the inclusion of other metadata in the indigenous language should not be thought only for the repository of this institution, but rather for other public institutions or other types of repositories. This allows the representation of indigenous ethnicities in scientific and academic production to be widely disseminated in all areas of knowledge, thus implementing cognitive and authorial justice, as well as putting the Law into practice.

All implementations made in the test environment will allow the UFSCar repository to truly be thought of as an instrument for archiving and accessing scientific information. It is also a means of preservation, visibility, appreciation, respect, and strengthening of languages and indigenous knowledge that are currently under threat.

Acknowledgements

To the National Council for Scientific and Technological Development (CNPq) for granting scholarships. To the indigenous students of UFSCar, for providing information that validated the development of the research."

References

Adler, Melissa. 2016. "The Case for Taxonomic Reparations". *Knowledge Organization*, 43, no. 8: 630-40. https://doi.org/doi.org/10.5771/0943-7444-2016-8-630

Alves, Lizandra de Sousa Santos, Roberta Cristina Dal'Evedove and Mariângela Spotti Lopes Fujita. 2022. "Avaliação do uso de vocabulário controlado em repositórios institucionais". *Informação@Profissões* 11, no. 1: 52–77. https://doi.org/10.5433/2317-4390.2022v11n1p52

English Name of Language	ISO 639-2 Code	Definition
Apurinã	apu	Apurinã is a Southern Maipurean language spoken in Amazonia by the Apurinã people.
Arapaso	arj	-
Bakairí	bkq	Bakairí (Bacairí) is a Cariban language of Brazil.
Baniwa	bwi	Baniwa (Baniva), or Baniwa from Içana (Baniua do Içana), is an Arawakan language spoken in Amazonas, Brazil, and in Venezuela.
Baré	bae	Baré (Barawana) is an Arawakan language, probably extinct, from Venezuela and Brazil. Aikhenvald (1999) reports "just a few old speakers left" of proper Baré, and that the Guinau variety was extinct. Ethnologue (2009) reports "no known speakers". Kaufman (1994) considers Baré proper, Guinau, and extinct Marawá to be distinct languages; Aikhenvald considers them dialects of a single languages. (Marawá is not the same language as Marawán).
Canela	ram	Canela is a Ge language spoken in Brazil.
Cocama language	cod	Cocama (Kokáma) is an indigenous language spoken by thousands of native people in western South America. It is spoken along the banks of the Northeastern lower Ucayali, lower Marañón, and Huallaga rivers and in neighboring areas of Brazil and an isolated area in Colombia. There are three dialects. The robust dialect is known as Cocama, Kokama, Ucayali, Xibitaoan, Huallaga, Pampadeque, and Pandequebo.
Cubeo	cub	The Cubeo language (otherwise known as Cuveo, Hehenawa, Kobeua, Kobewa, Kubwa, or Pamiwa) is a SOV language spoken by the Cubeo people and is a member of the central branch of the Tukano language. It has many lexical loans from the Nadahup languages and has a grammar which was apparently influenced by Arawak. The language is spoken in the Vaupés department, Cuduyari, and Querarií rivers and tributaries of Colombia. It is also spoken in Brazil.
Guajajára	gub	-
Guanano	gvc	Guanano (also Wanano) is a Tucanoan language spoken in the northwest part of Amazonas in Brazil and in Vaupés in Colombia.
Guarani	grn	-
Guarequena	gae	Guarequena (Warekena) is an Arawakan language of Brazil and Venezuela. It is one of several languages which goes by the generic name Baré.
Irantxe	irn	Irantxe (Iranxe, Iranshe), also known as Münkü (Mỹky), is an indigenous American language that is spoken in Mato Grosso, Brazil by about 200 people.
Kadiwéu	kbc	Kadiweu is a Mataco–Guaicuru language spoken by 1,200-1,800 people in Brazil. It is mainly a subject–verb–object language and its ISO 639-3 code is kbc.
Kaingang	kgp	The Kaingang language (also spelled Kaingáng) is an indigenous language belonging to the Gê language family spoken in the South of Brazil. The Kaingang nation has about 30,000 people, and about from 60% to 65% speak the language. The majority also speaks Portuguese.
Kayabí	kyz	Kayabí (Caiabi) is a Tupian language of Matto Grosso, Brazil.
Kayapó	txu	Kayapó, also known as Mẽbengokre, is a Ge language of Brazil. The majority are monolingual, and most who are bilingual speak other indigenous languages; perhaps 1% speak Portuguese.
Kamayurá	kay	The Kamayurá language (Kamaiurá in Portuguese) belongs to the Tupi–Guarani family and is spoken by the Kamayurá people of Brazil, who numbered about 290 individuals in 2004.
Kambiwá	xbw	Kambiwá Cambioá is an extinct unclassified language of Brazil. A couple dozen words were collected from two people in the 1960s, but by that time no-one spoke the language any longer. Apart from two apparent borrowings, none of the words are relatable to known languages.
Krenak	kqq	The Krenak language, or Botocudo, is the moribund sole surviving language of a small family believed to be part of the Macro-Gê languages. It was once spoken by the Botocudo people in Mato Grosso, but is now only known by older women.
Macuna	myy	-
Macushi	mbc	Macushi is the most populous of the Cariban languages, spoken by 30,000 in Brazil and Guyana. It is also spelled Makushi, Makusi, Makuxi, Macusi, Macussi, and also known as Teweya (Teueia).
Marúbo	mzr	Marúbo is a Panoan language of Brazil.

Table 3. Indigenous Language identifiers *(to be continued)*.
Source: Prepared by the authors from Library of Congress and Vocabs ROSSIO (2024).

English Name of Language	ISO 639-2 Code	Definition
Matsés	mcf	The Matsés language (sometimes called Mayoruna) is an indigenous language of the Peruvian and Brazilian Amazon basin which belongs to the Panoan language family.
Nhengatu	yrl	The Nheengatu language (in original Tupi pronunciation), often spelled Nhengatu, is an Amerindian language of the Tupi–Guarani family. It is also known by the Portuguese names língua geral da Amazônia and língua geral amazônica, both meaning "Amazonian General Language," or even by the Latin lingua brasilica (Brazilian Language). Nheengatu originated in northern Brazil in the 17th century as a lingua franca. Now known as nheengatu (also nhengatu, nyengatú, língua geral, geral, yeral), it is still spoken along the Rio Negro in northern Brazil (as well as in neighboring Colombia and Venezuela).
Pataxó language	pth	Pataxó or Pataxó Hã-Ha-Hãe is an extinct native language of Brazil formally spoken by the Pataxó people of the Bahía region and Minas Gerais, Pôsto Paraguassu in Itabuna municipality. It is unclassified. The 2,950 individuals in the Pataxó tribe now speak Portuguese. Pataxó Hã-Ha-Hãe was also known as Patashó, Pataxi, and Pataxó-Hãhaãi.
Pemon	aoc	The Pemon language is a Native American language of the Cariban family spoken by some 30,000 Pemon people, in Venezuela's Southeast, particularly in the Canaima National Park, in the Roraima State of Brazil and in Guyana.
Piratapuyo	pir	-
Rikbaktsa	rkb	The Rikbaktsa language, also spelled Aripaktsa, Erikbatsa, Erikpatsa and known ambiguously as Canoeiro, is a language spoken by the Rikbaktsa people of the Mato Grosso, Brazil, that forms its own branch in the Macro-Gê languages.
South American Indian languages	**sai**	-
Suruí	sru	Suruí, Paíter or Suruí-Paíter, is a Tupian language of Brazil.
Tariana	tae	Tariana (also Tariano) is an endangered Maipurean language spoken along the Vaupés River in Amazonas, Brazil by approximately 100 people. Another approximately 1500 people in the upper and middle Vaupés River area identify themselves as ethnic Tariana but no longer speak the language.
Tembé	tqb	-
Tereno (Terêna language)	ter	Terêna or Etelena is spoken by 15,000 Brazilians. The language has a dictionary and written grammar. Many Terênan people have low Portuguese proficiency. It is spoken in Mato Grosso do Sul. 20% are literate in their language, 80% literate in Portuguese.
Ticuna	tca	Tïcuna, or Tikuna, is a language spoken by approximately 40,000 people in Brazil, Peru, and Colombia. It is the native language of the Tïcuna people. Tïcuna is generally classified as an isolated language, but may be related to the extinct Yuri language. (See Tïcuna-Yuri.) It is a tonal language, and therefore the meaning of words with the same phonemes can vary greatly simply by changing the tone used to pronounce them.
Tucano	tuo	Tucano (also Tukana, Tucana, Tukano, Dasea, Jurutí, Juritί, Yurutí, Tariana, Tariano, Koneá, Koreá, Patsoka, Wahyara; autonym: Dahseyé) is a Tucanoan language spoken in Amazonas, Brazil, and Colombia.
Tupi languages	tup	Tupiniquim (Tupinaki) is a language which was spoken by Tupiniquim tribesmen in the Brazilian states of Espírito Santo and Bahia, and belonged to the Tupi–Guarani language family. It is now extinct. Its former speakers have switched to Portuguese. Alternative terms for the concept: Acatepec Tlapanec; Me'phaa, Acatepec; Tlapanec, Acatepec; Tupiniquim language.
Tuyuca	tue	Tuyuca (also Dochkafuara, Tejuca, Tuyuka, Dojkapuara, Doxká-Poárá, Doka-Poara, or Tuiuca) is an Eastern Tucanoan language (similar to Tucano) spoken by the Tuyuca people. The Tuyuca are an indigenous ethnic group of some 500-1000 people who inhabit the watershed of the Papuri, Inambú, and Tiquié rivers in the Colombian department of Vaupés and the Brazilian state of Amazonas.
Xakriabá	xkr	Xakriabá (also written Chakriaba, Chikriaba, Shacriaba) is an extinct Ge language formerly spoken in Minas Gerais, Brazil by the Xakriabá people, who today speak Portuguese.
Xavánte	xav	The Xavante language is a Ge language spoken by the Xavante people in about 170 villages in the area surrounding Eastern Mato Grosso, Brazil. The Xavante language is unusual in its phonology, its object–subject–verb word order, and its use of honorary and endearment terms in its morphology.

Table 3 continued

English Name of Language	ISO 639-2 Code	Definition
Wapishana	wap	Wapishana (Wapixana) is an Arawakan language of Guyana and Brazil.
Wasu	wsu	Wasu (Waçu, Wassú) is an extinct unclassified language of Brazil. The ethnic population is about 1,500.

Table 3 continued

dc.identifier.citation	WAWEMRA, Eudócio Tserewiwe. Relatos e narrativas A'uwe Xavante em diálogo com a Educação Especial. 2021. Trabalho de Conclusão de Curso (Graduação em Educação Especial) – Universidade Federal de São Carlos, São Carlos, 2021. Disponível em: https://repositorio.ufscar.br/handle/ufscar/15784.	*
dc.identifier.uri	https://repositorio.ufscar.br/handle/ufscar/15784	
dc.description.abstract	Este trabajo comienza con mi historia de vida en Aldeia Sangradouro, MT, del pueblo A'uwe Xavante, hasta llegar a la Universidad Federal de São Carlos, como estudiante de la Licenciado en Educación Especial. Basado en artículos sobre la historia y la educación de mi gente, publicado por familiares, escribo sobre nuestra nuestra historia, nuestro pueblo, la Escuela São José de Sangradouro y sobre Educación Especial. Con el desarrollo de este trabajo Trato de conocer cómo viven las personas con discapacidad en el pueblo de Sangradouro y reflexionar sobre la posibilidad de ayudar en la integración social de estas personas. A través de conversaciones informales. con familiares y personas con discapacidad, les presento algunas de sus historias y cómo viven en el pueblo. Teniendo también una hija diagnosticada con Discapacidad Intelectual, describo en profundidad de las relaciones familiares, actividades y desarrollo. a partir de una plan desarrollado para mi hija, describo la realización de actividades que involucran historias, diálogos, rutinas y convivencia con la familia y en el pueblo.	spa
dc.description.abstract	Īró ãmahã Sangradouro Rówatsu 'u inarada iréhã īhoimanadzémnahã ãhã rómhurihã, MT, īnhima a'wẽ remhã A 'wẽ Uptabi Universidade Federal de São Carlos, u īwitsidzémuhã, niha niha rómnhõré 'wana īhoimanahã curso de Licenciatura em Educação Especial remhã. Ītsiré a 'uwẽ te ãma irópé Rówatsu u rómreme iwẽtsihã duré rómnhõrénahã īte a 'uwẽ nahã, watsihoto wate rówatsu 'u, waró ãmahã, rómnhõrédzé São José de Sangradouro duré Educação Especial. Rómhuri itsa ẽtẽna warîni tewaihu 'uda eniha terehoimanadzara Ipótówatsédé daróbremhã Sangradouro duré róbdzanhamri iwaihu 'upe da ãma datsime dahoimanada i aho 're ãhã niwamnõrī. Róbdzanhamri daporepu 'u ãna ītsiré a 'uwẽ ma duré niwamnõrī ipódó aimawi wahoiré tsuruna ite rówatsu 'u duré eniha terehoimanadzara daróbremhã. Duré īra iréwa date iwaihu i 'rã iboi õ, wawatsu uwẽ ihoimanadzé tinanõrīmehã, watsihoto rómhuri tenatsi i amanharī dzé rówatsu udzéma, róbdzanhamri, ihoimanadzé duré itsimirómhuri titsitsãnawã nõrī mehã daróbremhã.	sai

Figure 5. Inclusion of other metadata in the indigenous language.
Source: UFSCar Institutional Repository – internal pre-test interface.

Andersen, Jack and Laura Skouvig, eds. 2017. *The Organization of Knowledge*: Caught Between Global Structures and Local Meaning. Bingley: Emerald Publishing.

Bazilio, A. P. M. 2022. "Repositório institucionais a serviço da Ciência, Tecnologia e Sociedade: uma análise sobre o RiUFF e LUME". PhD diss., Universidade Federal de São Carlos. https://repositorio.ufscar.br/handle/ufscar/17271

Brasil. 1988. *Constituição da República Federativa do Brasil de 1988*. http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm

Brasil. 2012. *Lei nº 12.711, de 29 de agosto de 2012*. https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/l12711.htm

CONARQ (Conselho Nacional de Arquivos). 2023. *Diretrizes para implementação de repositórios arquivísticos digitais confiáveis (RDC-Arq)* (2. versão). https://bdan.an.

dc.language.iso	por	por
dc.publisher	Universidade Federal de São Carlos	por
dc.rights	Attribution-NonCommercial-NoDerivs 3.0 Brazil	*
dc.rights.uri	http://creativecommons.org/licenses/by-nc-nd/3.0/br/	*
dc.subject	Educação Indígena A'uwe Xavante	por
dc.subject	Educação Especial	por
dc.subject	História de vida	por
dc.subject	Rómnhõré A 'wẽ Uptabi	sai
dc.subject	Rómnhõré itsipe	sai
dc.subject	Rówatsu 'u Dahoimanadzépte	sai
dc.title	Relatos e narrativas A'uwe Xavante em diálogo com a Educação Especial	por
dc.title.alternative	Informes y narrativas A'uwe Xavante en diálogo con la Educación Especial	spa
dc.title.alternative	Rówatsu 'u duré iwatsu 'u A'uwe Xavante róbdzanhamri rómnhõré îpe tsiré	sai
dc.type	TCC	por

Figure 6. Inclusion of other metadata in the indigenous language.
Source: UFSCar Institutional Repository – internal pre-test interface.

gov.br/server/api/core/bitstreams/2436b88e-b5c7-4393-baab-ab6c933fed82/content

(CNMP) Conselho Nacional do Ministério Público. 2024. *Sistema ELO: Plataforma de gestão eleitoral*. https://elo.cnmp.mp.br/login.seam

(CNMP) Conselho Nacional do Ministério Público and (CNJ) Conselho Nacional de Justiça. 2012. *Resolução Conjunta nº 3, de 19 de abril de 2012*. https://atos.cnj.jus.br/files//resolucao_conjunta/resolucao_conjunta_3_19042012_25042013155653.pdf

Daltro, Mônica Ramos and Anna Amélia de Faria. 2019. "Relato de experiência: Uma narrativa científica na pós-modernidade". *Estudos e Pesquisas em Psicologia* 19, no. 1: 223-37. http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1808-42812019000100013&lng=pt&tlng=pt

Doyle, A. M. 2013. "Naming, Claiming, and (Re)creating: Indigenous Knowledge Organization at the Cultural Interface". PhD diss., University of British Columbia.

Dspace. n.d. *About DSpace*. Dspace (website). https://dspace.lyrasis.org/about/

Freitas, Marina Penteado de. 2019. "Autoarquivamento e representação de assunto: Estudo analítico de teses e dissertações do Repositório Institucional da UFSCar". PhD diss., Universidade Federal de São Carlos. https://repositorio.ufscar.br/handle/ufscar/11850

Fujita, Mariângela Spotti Lopes and Jessica Cristina Panuto. 2024. "Guidelines on Assigning The Subjects of Theses and Dissertations in Repositories". *IFLA Journal* 50, no. 1: 160-9. https://doi.org/10.1177/03400352231217275

Fujita, Mariângela Spotti Lopes. 2022. "Sistematização de modelo de avaliação do controle de vocabulários em repositórios". *Revista Digital de Biblioteconomia e Ciência da Informação* 20. https://doi.org/10.20396/rdbci.v20i00.8668751

Fujita, Mariângela Spotti Lopes, Carmen Agustin-Lacruz, Jéssica Beatriz Tolare, Ana Lúcia Terra and Germa Bueno de la Fuente. 2023. "Institutional Repositories and Knowledge Organization: A Bibliographic Study From Library and Information Science". *Education for Information* 39, no. 1: 51-66. http://hdl.handle.net/11449/250038

Fujita, Mariângela Spotti Lopes, Roberta Cristina Dal'Vedove and Jéssica Cristina Panuto. 2024. "The Indexing Policy in the Practices of Brazilian Institutional Repositories: A Diagnostic Study From the Perception of Managers and Indexers". *Journal of Librarianship and Information Science* 56, no. 2: 1-19. https://doi.org/10.1177/09610006221145544

Figure 7. Home page of an indigenous authored item.
Source: UFSCar Institutional Repository – internal pre-test interface.

Fujita, Mariângela Spotti Lopes, Rosane Rodrigues de Barros Ribas, Milena Maria Scarpino Rodrigues and Telma Jaqueline Dias Silveira. 2022. "Metadados de assuntos de teses e dissertações em repositórios: estudo exploratório sobre controle de vocabulário". *Em Questão* 28, no. 4: 1-27. https://doi.org/10.19132/1808-5245284.123710

Gilliland, Anne J. 2012. "Contemplating Co-creator Rights in Archival Description". *Knowledge Organization*, 39, no. 5: 340-6. https://doi.org/doi.org/10.5771/0943-7444-2012-5-340

Gonçalez, P. R. V. 2017. "Repositórios arquivísticos digitais confiáveis: identificação de requisitos com ênfase no acesso à informação". PhD diss., Universidade Estadual Paulista.

Gracioso, Luciana de Souza, Luzia Sigoli Fernandes Costa and Omar Lopes Silva. 2023. "Desafios da Organização do Conhecimento para a Representação Temática dos Saberes Originários em Sistemas de Informação: Reflexões Preliminares". In *Organização e Representação do Conhecimento em Diferentes Contextos: Desafios e Perspectivas na Era da Datificação*, edited by Natália Bolfarini Tognoli, Ana Cristina de Albuquerque e Brígida Maria Nogueira Cervantes, 615-21. Londrina: ISKO-Brasil.

Green, Rebecca. 2015. "Indigenous Peoples in the U.S., Sovereign Nations, and the DDC". *Knowledge Organization* 42, no. 4: 211-21. https://doi.org/10.5771/0943-7444-2015-4-211

IBGE (Instituto Brasileiro de Geografia e Estatística). 2010. *Censo*. Rio de Janeiro: IBGE. https://censo2010.ibge.gov.br/

ILO (International Labour Organization). 1989. Indigenous and Tribal Peoples Convention, 1989 (No. 169).

Geneva: International Labour Organization. https://www.ilo.org/dyn/normlex/en/f?p=1000:12100:0::NO:12100:P12100_INSTRUMENT_ID:312314

IPOL (Instituto de Investigação e Desenvolvimento em Política Linguística). 2017. Línguas cooficializadas nos municípios brasileiros. Florianópolis: IPOL. http://ipol.org.br/tag/linguas-indigenas-2/

Lee, Tamara, Sarah Dupont and Julia Bullard. 2021. "Comparing the Cataloguing of Indigenous Scholarships: First Steps and Findings". *Knowledge Organization* 48, no. 4: 298-306. https://doi.org/10.5771/0943-7444-2021-4-298

Leite, Fernando César Lima. 2009. *Como gerenciar e ampliar a visibilidade da informação científica brasileira: Repositórios institucionais de acesso aberto*. Brasília: Ibict.

Library of Congress. 2024. "Codes for the Representation of Names of Languages". *ISO 639-2* (website). https://www.loc.gov/standards/iso639-2/php/English_list.php

Littletree, Sandra, Miranda Belarde-Lewis and Marisa Duarte. 2020. "Centering Relationality: a Conceptual Model to Advance Indigenous Knowledge Organization Practices". *Knowledge Organization* 47, no. 5: 410-26. https://doi.org/10.5771/0943-7444-2020-5-410

Luciano, Gersem and Wager Amaral. 2021. "Povos Indígenas e Educação Superior no Brasil e no Paraná: Desafios e Perspectivas". *Integración y Conocimiento* 10, no. 2: 13–37. https://doi.org/10.61203/2347-0658.v10.n2.34069

Periotto, Caroline, Felipe Augusto Arakaki, Luciana de Souza Gracioso, Luzia Sigoli Fernandes Costa and Jair de Jesus Lages Massa. 2024. "Representatividade de etnia indígena em produção científica no Repositório Institucional da UFSCar: reflexões e ações preliminares". *VII Workshop De Informação, Dados E Tecnologia – WIDaT 2024* 7. https://doi.org/10.22477/vii.widat.171.

Sistema Integrado de Bibliotecas da UFSCar. nd. "RI UFSCar". *Sistema Integrado de Bibliotecas da UFSCar* (website). https://www.sibi.ufscar.br/repositorio

Sundström, Admeire da Silva Santos. 2023. "Ethical Issues from Decolonial Practices in Knowledge Organization : The Case of Indigenous Collections in Världskulturmuseet". *Knowledge Organization,* 50, no. 6, 439–51. https://doi.org/10.5771/0943-7444-2023-6-439

Torino, Emanuelle. 2017. "Políticas em repositórios digitais: das diretrizes à implementação". In *Repositórios digitais: teoria e prática* edited by Fernando Luiz Vechiato, Clediane de Araujo Guedes, Débora Costa Araújo Di Giciaomo, Elisângela Alves de Moura, Emanuelle Torino Maria Aniolly Queiroz Maia and Tércia Maria Souza de Moura Marques, 91-114. Curitiba: EDUTFPR.

Tribunal Superior Eleitoral. Secretaria de Gestão da Informação e do Conhecimento. 2021. *Resolução TSE nº 23.659/2021: Dispõe sobre a gestão do Cadastro Eleitoral e sobre os serviços eleitorais que lhe são correlatos.* Tribunal Superior Eleitoral. https://www.tse.jus.br/legislacao/compilada/res/2021/resolucao-no-23-659-de-26-de-outubro-de-2021

Turner, Hannah. 2017. "Organizing Knowledge in Museums: A Review of Concepts and Concerns". *Knowledge Organization*, 44, no. 7: 472-84. https://doi.org/10.5771/0943-7444-2017-7-472

Universidade Federal de São Carlos. 2016. *Resolução ConsUni n° 835, de 04 de março de 2016. Dispõe sobre a criação, a política de implantação e a governança do Repositório Institucional da UFSCar*. https://repositorio.ufscar.br/static/politica-do-ri-ufscar.pdf

Universidade Federal de São Carlos. 2024. *Resolução CO/SIBI nº 3, de 22 de fevereiro de 2024. Aprova a Política do Repositório Institucional da UFSCar*. https://www.sibi.ufscar.br/arquivos/resolucao-cosibi-no-1.pdf

Vidotti, Silvana Aparecida Borsetti Gregorio, Emanuelle Torino and Caio Saraiva Coneglian. 2021. "#SejaJUSTOeCUIDADOSO: princípios FAIR e CARE na gestão de dados de pesquisa". In *Princípios FAIR aplicados à gestão de dados de pesquisa,* edited by Luana Farias Sales, Viviane Santos de Oliveira Veiga, Patrícia Henning and Luís Fernando, 201-14. Rio de Janeiro: IBICT.

Villanueva, Cristina B. 2022. "Language Subject Access to Indigenous Materials: The Case of the Philippine Cordillera". *Cataloguing and Classification Quarterly*, 60, no. 3-4: 297-314. https://doi.org/10.1080/01639374.2022.2075512

Provocations of Process in Critical Knowledge Organization Work

Julia Bullard

University of British Columbia, 470-1961 East Mall,
Vancouver, BC Canada V6T 1Z1
julia.bullard@ubc.ca

Julia Bullard is an assistant professor at the iSchool and teaches courses in knowledge organization, classification theory, and information systems. Her research looks at the way that knowledge organization systems are designed, implemented, revised, broken or fixed. She is interested in the design of organization systems and how these systems might align with the values and activities of their communities. Her dissertation work focused on classification systems – the designed structure of terms, their relationships and the rules for applying these terms tor documents or people. Julia Bullard has an MA in Cultural Studies and Critical Theory (McMaster), an MLIS from UBC iSchool, and a Ph.D. from the Information School at the University of Texas at Austin. She has published extensively in classification and knowledge organization systems, and the design of organizing systems such as metadata schemas and controlled vocabularies.

Bullard, Julia. 2024. "Provocations of Process in Critical Knowledge Organization Work". *Knowledge Organization* 51, no. 8: 660-666. 49 references. DOI:10.5771/0943-7444-2024-8-660.

Abstract: In this paper, I argue that the most provocative work in critical knowledge organization is happening at the level of *process*. I present three persistent assumptions about knowledge organization work and current provocations that challenge them. First, that systems should be seamless and not reveal the work behind them. Second, that systems should achieve a single authorial voice through consistency, precedent, and patterns. Third, that knowledge organization systems are best applied with minimal interpretation on the part of the worker. The provocations against each of these assumptions come from current and highly regarded work in the field, indicating greater respect and visibility for the processes behind knowledge organization systems.

Received: 10 July 2024; Accepted 25 July.

Keywords: critical knowledge organization; knowledge organization systems; authorial voice; process.

1.0 Introduction

The common image of critical knowledge organization is reparative or liberatory work achieved through more equitable, authentic, and respectful terminology and term relationships. In this paper, I argue that the most provocative work in critical knowledge organization is happening at the level of *process*. Knowledge organization, as infrastructure, tends to be invisible until it does not work, meaning that those most unjustly treated by colonial, cisheteropatriarchal worldviews are positioned to notice the operations of a controlled vocabulary or classification while those benefitting from those power relations do not (Star 1999; Huvila 2009; Lee 2015; Bullard et al. 2022). This essay engages in a second, deeper layer of invisibility: the design work that goes into the system itself. These two types of invisibility are entangled; a seamless experience of information retrieval makes it easier to assume that the system is fully automated *and* ignorance of the design and maintenance labour of the system makes it easier for users to assume the system is objective or inevitable.

2.0 Literature Review

Critical approaches are fundamentally about questioning the assumptions we treat as given and applying a power analysis to the persistence of those assumptions. In knowledge organization, common assumptions amenable to a critical analysis are that our dominant systems are accurate, objective, inevitable, and innocent. Critical knowledge organization work is often about identifying and exposing the harms and distortions of these existing systems. Berman (1971) provided an extensive review of such issues in the *Library of Congress Subject Headings* (LCSH) and Olson and Schlegl (2001) provided a meta-analysis of further critiques of subject access through to 1999. As description and documenta-

tion, critical knowledge organization identifies points of failure and the power relations that are reified and exacerbated by our systems, often identifying colonialism (Pacey 1989; Dudley 2017; Biswas 2018; Bullard et al. 2022), racism (Howard and Knowlton 2018; Noble 2018; Baron and Gross 2021), sexism (Olson 2002), and heterosexism and transphobia (Christensen 2008; Howard and Knowlton 2018; Watson 2020; Henry et al. 2022) as the dominant logics reproduced by subject access systems in particular.

Taking such critiques as a point of departure, a productive realm for critical knowledge organization studies is the repair of such systems, remedying what was identified through the descriptive approach. Scholars such as Berman and Olson have typically done both and paired their analysis with proposals for repair and intervention into dominant systems including Library of Congress Subject Headings (Olson 2000; Berman and Gross 2017) and the Dewey Decimal System (Olson 1998). Direct critiques of current standards can intervene in common practice (Billey et al. 2014; Billey and Drabinski 2017), impacting practices at hundreds or thousands of libraries and cultural heritage institutions following such guidelines. Alongside these efforts to ameliorate the harm of discriminatory systems, critical knowledge organizations scholars propose challenges to a "politics of correction," advocating for instruction and dialogue (Drabinski 2013) and admitting the limitations of subject-based access for identity (Stahl 2024).

An alternate approach to repair is liberation from reliance on centralized and colonial systems, producing community-centred systems that prioritize practices and cultural perspective otherwise marginalized (Allard et al. 2021; Berg et al. 2022), such as Indigenous ways of knowing (Doyle et al. 2015; Littletree and Metoyer 2015) and queer identities (Bullard et al. 2020; Homosaurus 2020). These liberatory projects often champion a departure from universality, embracing a minority or multiplicity of viewpoints with a clear set of values and priorities, aligning to a contemporary model of knowledge organization in which bias is not denied but identified (Feinberg 2007). These liberatory projects are often built alongside and within a particular collection, such as in the systems built for archiveofourown.org (Price 2019) but may also be networked and interoperable with distributed and diverse collections (Homosaurus 2020).

Both the repair and liberation frameworks of critical knowledge organization engage with practical and present use of systems. Distinct from general approaches to assessing and improving the quality of knowledge organization infrastructure (Snow 2017), the critical approaches address how the (dys)function of such systems are not evenly distributed, such that they are invisible (Star and Strauss 1999) and provide a seamless user experience for some users while those made marginal experience friction and alienation. While such work aims to produce outcomes in the form of more just classes, terms, and term relationships, they often engage in criticisms of processes, power, and epistemology, linking the harms of the status quo to the histories and commitments of dominant systems. A few such works connect the outward harms of the systems with the internal structures of labour, noting that structures of automation, precarity, outsourcing, and reliance on vendor metadata are inhospitable to the types of attention and accountability required for repair or liberatory practices.

The remainder of this paper will expand on the overlap between critical knowledge organization concerned about knowledge organization systems at they exist in the world (as schedules, term lists, or schemas) and critical knowledge organization about the processes that produce the system as it exists.

3.0 Assumptions and provocations on process

Throughout this section, I introduce critical knowledge organization processes through the assumptions they challenge. Many of these assumptions are so ingrained into dominant knowledge organization practice that they can be perpetuated without being explicitly stated, and as such it is easier to identify these principles through criticisms or provocations against them.

3.1 Seamlessness

A common goal of user experience design – broader than but applicable to knowledge organization design – is the achievement of a "seamless" experience for the user. Seamless systems are those in which the user experiences no friction, unintuitive steps, or awkward workarounds in their operation of the product. The metaphor draws from physical construction methods (sewing, soldering) in which skillful joining of multiple pieces will leave little evidence of that intervention in the material.

As with the subtle seam of joined fabric or metal, seamlessness is an illusion or a quality of perception rather than a characteristic of the system itself. Behind any system approaching this status is an un(der)acknowledged set of practices stitching together different and changing data sources, standards, and technologies, known variously as articulation work (Schmidt and Bannon 1992) or, when the illusion is of an automated system, heteromation (Ekbia and Nardi 2014). Scholars of articulation work and heteromation point to the enormous of labour that goes into making end user experiences seamless and, especially for heteromation, to keeping the user from perceiving that very labour.

The work of the *Unseen Labor* embroidery project is a particularly apt provocation against the idea of seamlessness. Sourced from an engaged community of catalogue and

metadata workers in the United States, Canada, and the United Kingdom, Unseen Labor is an embroidery collection exemplifying the creativity, beauty, and humour of their behind-the-scenes work (Kardos 2022; 2023). Contrasting against seamlessness, in which the stitching vanishes within the joined fabric, embroidery is an artwork of visible stitches. Projects like *Unseen Labor* flip the background and foreground of library work, putting the detail-oriented and otherwise invisible work of cataloguing into exhibit boxes and displays.

3.2 Coherent authorial voice

Arguing that a knowledge organization system has a voice was once itself a provocation (Feinberg 2007), posing the idea that systems had something to communicate as well as a function to serve in the retrieval of objects. Feinberg's approach was provocative in that each information system in her analysis had a notably coherent point of view, making the idea of voice recognizable through a critical textual analysis method. Her point is applicable to knowledge organization systems broadly: systems are communicative devices alongside and through their labelling, categorization, and retrieval functions.

Applying Feinberg's concept beyond the exemplars in her analysis, we can recognize the emergence of an authorial voice in other knowledge organization systems. She notes that authorial voice as a quality of a text rather than as an outcome of authorial intent was already well-established in literary theory; a text having multiple and distributed authors is not incompatible with an authorial voice emerging in the text itself. The intent or the mechanisms to establish a coherent voice can be traced to the standardization and consistency guidelines of core knowledge organization systems. The Library of Congress's Subject Headings Manual facilitates this coherence across decades of expansion and revision and through the thousands of cataloguers who submit subject heading proposals through key principles of evaluation such as:

- Does the proposed heading employ neutral (i.e., unbiased) terminology?
- Does the proposed heading conform to patterns and precedents in LCSH with respect to wording, form, and style, and to guidelines provided in the SHM? (Library of Congress 2020)

as well as through the centralized editorial process in the Washington, D.C. office. Consistency through patterns and precedents has practical outcomes: users who develop literacy in LCSH formats will be empowered to conduct efficient and effective searches when they can predict and extrapolate to new subject searches. Among the 11 questions presented in this section of the LCSH manual, the consecutive sequence of two above is meaningful; through a Kantian framework (Fox and Reece 2012), the Library of Congress's approach to ethical and effective organization of resources is to endeavour to treat every subject identically, neutrally, and in mirrored and parallel styles across the entire vocabulary.

In 2023, a working group of the Subject Analysis Committee of the American Library Association submitted (ALA) a report to the Library of Congress, recommending that the institution end centralized control in favour of external oversight of LCSH and its other vocabularies (Badger et al. 2023). Drawing from a wide selection of controlled vocabularies and their models of editorial processes, the report advocates for an editorial board composed of groups described, subject experts, and international users, as well as incorporating input from publishers, editors, non-cataloguers, and non-library catalogue users. The provocativeness of this proposal is tempered by the context the report provides on its consistency with other, positive changes underway at the institution as well as the established precedent in other extant controlled vocabularies.

Editorial boards, like those of the Homosaurus the report uses as an exemplar, pose a model of multivocal rather than singular and coherent authorial voice. Contrasting to the categorical imperative identified by Fox and Reece as being typical of standardized, outcome-independent processes of knowledge organization system design, processes like Homosaurus's are more likely to enact an ethic of care in which terms and term relationships are created and evaluated in their contexts of historical and ongoing oppression and known harms (Fox and Reece 2012). This multivocality can be seen in decisions such as Homosaurus version 3.1 (December 2021) establishing the noun order of "LGBTQ+ [term]" for terms like "LGBTQ+ parents" and version 3.2 (June 2022) making batch changes to 85 of those 259 terms, those involving race and ethnicity, to the reverse order: "Afro-Canadian LGBTQ+ people." The history note explains: "To acknowledge their multiple marginalization, terms for people of color who are LGBTQ+ use the following format: Racial or ethnic identity followed by gender or sexual identity (e.g., Black lesbians)" (Homosaurus 2020).

Editorial processes which distribute the authority of system expansion and revision decisions are more likely to produce such multivocal and context-aware policies and outcomes. Outside of traditional librarianship, a provocation against assumed necessity of centralized control has persisted for over 15 years in the form of archiveofourown.org's curated folksonomy (Archive of Our Own, n.d.; Bullard 2019; Price 2019). With all users able to add free-text tags to their uploaded works, and the site itself growing by millions of works a year, it takes a hundreds of hours a week of volunteer, specialist labour to provide constant maintenance of the back-end thesaurus structure, weaving new terms and

concepts into networks of preferred terms, variants, and term relationships (Bullard 2016; 2019; Price 2019). As with other core principles of the website's design (Fiesler et al. 2016), the primacy of creativity and multivocality are embedded in the process of the vocabulary's construction and maintenance.

3.3 Without interpretation

A final assumption of knowledge organization process is that they minimize interpretation. Principles like literary warrant, in which the terminology and proportions of a system are determined by the language, breadth, and specificity of the collection itself, set up processes in which the collection drives the system. Such processes allow knowledge organization workers (and their managers and institutions) to defer responsibility for biases in the system to the composition of the published work; asymmetries and pejorative terminology are not ontological statements made by the system's editors but rather epistemologies derived from works the system provides access to. This maneuver of ethical-sidestepping is less obvious than the labour implications of such processes. Avoiding interpretation may save some conceptual labour on the part of knowledge organization workers though it commits such systems to constant revision, the fractures of back-cataloguing older works with the terminology of current scholarship, and the surveillance and repair work of identifying and incorporating new concepts, trends, and phenomena as they appear in published works.

A recent provocation against minimal interpretation can be found in recent knowledge organization work from the Rare Books and Manuscripts Section (RBMS) of the ALA. A 2020-2022 working group focusing on updating the "Literature of prejudice" heading in the RBMS vocabulary aimed to facilitate the labeling of historical and contemporary primary materials that manifest prejudice and racism, both to support research on those phenomena and to provide context on materials through the catalogue record (Bychowski et al. 2023). The revised heading "Prejudicial works" now has an expanded set of narrower terms, with the following scope note: "Use for works that exhibit hostility toward or bias against a particular group or groups of people based on religion, race, ethnicity, gender, sexuality, ability, creed, national origin, etc. For works that express hatred or advocate harm toward a particular group of people, use: Hate works." (https://id.loc.gov/vocabulary/rbmscv/cv00108.html). Bychowski and co-authors note that the terms and their application to works is inherently at odds with "the privilege to do nothing," or the persistent but questioned maxim that cataloguing work should (or could) be neutral.

The existence of "Prejudical works" and its narrower terms questions the assumption that cataloguers apply work without judgment or interpretation. In order to apply terms such as these, cataloguers must interpret works not necessarily as they present themselves but in their context as actors in a struggle between human rights and persistent and harmful systems of oppression. Most starkly, these RBMS CV terms are in tension with the recently released *Cataloguing Code of Ethics* (Cataloging Ethics Steering Committee 2021) which includes the principle to "commit to describing resources without discrimination whilst respecting the privacy and preferences of their associated agents." This principle is most relevant to name authority work and promises to improve creator agency in representation and updating of names and the inclusion of personal information in name authority records. However, in describing resources, terms like "Prejudicial materials" can only be applied by violating the preferences of their associated agents in favour of other principles that seek to disrupt the instantiation of societal prejudices in cataloguing work.

4.0 Conclusion

The critical knowledge organization work I have presented here challenges three persistent assumptions in our field: that work should be seamless, have a coherent authorial voice, and should avoid interpretation. While some of the cited projects have their most apparent outcomes in the knowledge organization systems themselves – their classes, terms, and term relationships – the most radical outcomes can be traced back to radical departures from common processes and their long-held assumptions. In addition to producing more just, accurate, and usable systems, these changes in process also hold the potential to improve outcomes for knowledge organization workers themselves.

Skilled, intentional, and accountable knowledge organization work cannot compete with automated or machine learning approaches for scale. Human labour can instead differentiate itself by showing that process matters. These provocations share a sense of respect and recognition of the human labour of knowledge organization work. Should these approaches gain wider adoption in the field, it will add to workers' tools for making visible the value of their labour (Miksa 2008; Borie et al. 2015; Belantara and Drabinski 2022). The value of knowledge organization labour has long been difficult to assert, and processes that make that labour invisible (through seamlessness), centralized (through standardization and the imposition of a single authorial voice), and that denies the workers' perspective and judgement (through the denial of interpretive work) undermine workers' struggles for job security and pay equity.

Critical approaches are, by definition, challenges against the status quo and obsolescence is an inherent part of their success; what is provocative now, in 2024, may be an established mode of work in the future. Knowledge organization

work with its focuses on revision and maintenance accepts that systems must change; if the critical knowledge organization work described here gains traction, we will see that the way our systems change is itself subject to change.

References

Allard, Danielle, Shawna Ferris, Amy Lebovitch, Jenn Clamen and Micheline Hughes. 2021. "Ethical Accountability and High-Stakes Recordkeeping: Discussions from the Sex Work Activist Histories Project". In *Proceedings of the Annual Conference of CAIS / Actes Du Congrès Annuel de l'ACSI*. https://doi.org/10.29173/cais1200.

Archive of Our Own. n.d. "Tags FAQ." Accessed June 13, 2018. https://archiveofourown.org/faq/tags?language_id=en.

Badger, Allison, Allison Bailund, Julia Bullard, Jamie Carlstone, Crystal Casparis, Nicola Cook, Melissa De Fino *et al*. 2023. *Report of the SAC Working Group on External Review of Library of Congress Vocabularies*. Report. ALA Core Subject Analysis Committee. https://alair.ala.org/handle/11213/20012.

Baron, Jill E. and Tina Gross. 2021. "Sorrow, Fury, Helplessness, and Cynicism: An Account of the Library of Congress Subject Heading 'Illegal Aliens'". In *Borders and Belonging: Critical Examinations of Library Approaches toward Immigrants*, edited by Ana Ndumu. Sacramento, CA: Library Juice Press.

Belantara, Amanda and Emily Drabinski. 2022. "Working Knowledge: Catalogers and the Stories They Tell". *KULA: Knowledge Creation, Dissemination, and Preservation Studies* 6, no. 3: 1–10. https://doi.org/10.18357/kula.233

Berg, Magnus, Satwinder Kaur Bains and Sadhvi Suri. 2022. "The South Asian Canadian Digital Archive Thesaurus: Thesaurus Construction for the South Asian Canadian Diaspora." *KULA: Knowledge Creation, Dissemination, and Preservation Studies* 6, no. 3: 1–7. https://doi.org/10.18357/kula.223

Berman, Sanford. 1971. *Prejudices and Antipathies: A Tract on the LC Subject Heads Concerning People*. Metuchen, NJ: Scarecrow Press.

Berman, Sanford and Tina Gross. 2017. "Expand, Humanize, Simplify: An Interview with Sandy Berman". *Cataloging & Classification Quarterly* 55, no. 6: 347–60. https://doi.org/10.1080/01639374.2017.1327468.

Billey, Amber, and Emily Drabinski. 2017. "Cataloging, Gender, and RDA Rule 9.7". *American Library Association*, Webinar, March 15, 2017. http://www.ala.org/alcts/confevents/upcoming/webinar/031517.

Billey, Amber, Emily Drabinski and K. R. Roberto. 2014. "What's Gender Got to Do with It? A Critique of RDA 9.7". *Cataloging & Classification Quarterly* 52, no. 4: 412–21. https://doi.org/10.1080/01639374.2014.882465

Biswas, Paromita. 2018. "Rooted in the Past: Use of 'East Indians' in Library of Congress Subject Headings". *Cataloging & Classification Quarterly* 56, no. 1: 1–18. https://doi.org/10.1080/01639374.2017.1386253

Borie, Juliya, Kate MacDonald and Elisa Sze. 2015. "Asserting Catalogers' Place in the 'Value of Libraries' Conversation". *Cataloging & Classification Quarterly* 53, no. 3–4: 352–67. https://doi.org/10.1080/01639374.2015.1007543

Bullard, Julia. 2016. "Motivating Invisible Contributions: Framing Volunteer Classification Design in a Fanfiction Repository." In *GROUP '16: Proceedings of the 19th International Conference on Supporting Group Work 13-16 November Sanibel Island, FL, USA,* 181–93. New York: ACM Press. https://doi.org/10.1145/2957276.2957295.

Bullard, Julia. 2019. "Curated Folksonomies: Three Implementations of Structure through Human Judgment". *Knowledge Organization* 45, no. 8: 643–52. https://doi.org/10.5771/0943-7444-2018-8-643.

Bullard, Julia, Amber Dierking and Avi Grundner. 2020. "Centring LGBT2QIA+ Subjects in Knowledge Organization Systems". *Knowledge Organization* 47, no. 5: 393–403. https://doi.org/10.5771/0943-7444-2020-5-393

Bullard, Julia, Brian Watson and Caitlin Purdome. 2022. "Misrepresentation in the Surrogate: Author Critiques of 'Indians of North America' Subject Headings". *Cataloging & Classification Quarterly* 60, no. 6–7: 599–619. https://doi.org/10.1080/01639374.2022.2090039

Bychowski, Brenna, Ryan Hildebrand, Sarah Hoover and Lauren Reno. 2023. "Describing Special Collections: Two Projects to Modernize the RBMS Controlled Vocabularies." *Cataloging & Classification Quarterly* 61, no. 5–6: 626–45. https://doi.org/10.1080/01639374.2023.2193950

Cataloging Ethics Steering Committee. 2021. *The Cataloguing Code of Ethics*. https://cfla-fcab.ca/wp-content/uploads/2022/09/Cataloguing-Code-of-Ethics.pdf

Christensen, Ben. 2008. "Minoritization vs. Universalization: Lesbianism and Male Homosexuality in LCSH and LCC". *Knowledge Organization* 35, no. 4: 229–38. https://doi.org/10.5771/0943-7444-2008-4-229

Doyle, Ann Mary, Kimberley Lawson and Sarah Dupont. 2015. "Indigenization of Knowledge Organization at the Xwi7xwa Library". Journal of Library and Information Studies 13, no. 2: 107-34. https://doi.org/10.14288/1.0103204

Drabinski, Emily. 2013. "Queering the Catalog: Queer Theory and the Politics of Correction." *The Library Quar-*

terly 83, no. 2: 94–111. https://doi.org/10.1086/669547.

Dudley, Michael Q. 2017. "A Library Matter of Genocide: The Library of Congress and the Historiography of the Native American Holocaust". *International Indigenous Policy Journal* 8, no. 2. https://doi.org/10.18584/iipj.2017.8.2.9

Ekbia, Hamid R and Bonnie Nardi. 2014. "Heteromation and Its (Dis)Contents: The Invisible Division of Labor between Humans and Machines". *First Monday* 19, no. 6. https://doi.org/10.5210/fm.v19i6.5331

Feinberg, Melanie. 2007. "Hidden Bias to Responsible Bias: An Approach to Information Systems Based on Haraway's Situated Knowledges". *Information Research* 12, no. 4. http://informationr.net/ir/12-4/colis07.html

Fiesler, Casey, Shannon Morrison and Amy S Bruckman. 2016. "An Archive of Their Own: A Case Study of Feminist HCI and Values in Design". In *CHI '16: Proceedings of the ACM Conference on Human Factors in Computing Systems 7-12 May 2016, San Jose, CA, USA, 2574-85*. https://doi.org/10.1145/2858036.2858409.

Fox, Melodie J. and Austin Reece. 2012. "Which Ethics? Whose Morality?: An Analysis of Ethical Standards for Information Organization" *Knowledge Organization* 39, no. 5: 377–83. https://doi.org/10.5771/0943-7444-2012-5-377

Henry, Tiffany, Rhonda Kauffman and Anastasia Chiu. 2022. "The Old and the Prudish: An Examination of Sex, Sexuality, and Queerness in Library of Congress Classification – In the Library with the Lead Pipe." *In the Library with the Lead Pipe*, no. June. https://www.inthelibrarywiththeleadpipe.org/2022/the-old-and-the-prudish/.

Homosaurus. 2020. "Homosaurus: An International LGBTQ Linked Data Vocabulary". Homosaurus Vocabulary Site (website). http://homosaurus.org/.

Howard, Sara A. and Steven A. Knowlton. 2018. "Browsing through Bias: The Library of Congress Classification and Subject Headings for African American Studies and LGBTQIA Studies". *Library Trends* 67. no. 1: 74–88. https://doi.org/10.1353/lib.2018.0026

Huvila, Isto. 2009. "Ecological Framework of Information Interactions and Information Infrastructures". *Journal of Information Science* 35, no. 6: 695–708. https://doi.org/10.1177/0165551509336705

Kardos, Ann. 2022. *Unseen Labor*. University of Massachusetts Amherst Libraries.

Kardos, Ann. 2023. "Using Art to Talk and Think about Metadata: An Experiment". *College & Research Libraries News* 84, no. 3: 346-8. https://doi.org/10.5860/crln.84.9.346

Lee, Wan-Chen. 2015. "Culture and Classification: An Introduction to Thinking about Ethical Issues of Adopting Global Classification Standards to Local Environments". *Knowledge Organization* 42, no. 5: 302–7. https://doi.org/10.5771/0943-7444-2015-5-302

Library of Congress. 2020. H 204 Evaluating Subject Proposals. https://www.loc.gov/aba/publications/FreeSHM/H0204.pdf

Littletree, Sandra and Cheryl A. Metoyer. 2015. "Knowledge Organization from an Indigenous Perspective: The Mashantucket Pequot Thesaurus of American Indian Terminology Project". *Cataloging & Classification Quarterly* 53, no. 5–6: 640–57. https://doi.org/10.1080/01639374.2015.1010113

Miksa, Shawne D. 2008. "You Need My Metadata: Demonstrating the Value of Library Cataloging". *Journal of Library Metadata* 8, no. 1: 23–36. https://doi.org/10.1300/J517v08n01_03

Noble, SU. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.

Olson, Hope A. 1998. "Mapping beyond Dewey's Boundaries: Constructing Classificatory Space for Marginalized Knowledge Domains". *Library Trends* 47, no. 2: 233-54.

Olson, Hope A. 2000. "Difference, Culture and Change: The Untapped Potential of LCSH". *Cataloging & Classification Quarterly* 29, no. 1–2: 53–71. https://doi.org/10.1300/J104v29n01_04

Olson, Hope A. 2002. *The Power to Name: Locating the Limits of Subject Representation in Libraries*. Kluwer Academic Publishers. https://doi.org/10.1007/978-94-017-3435-6

Olson, Hope A. and Rose Schlegl. 2001. "Standardization, Objectivity, and User Focus: A Meta-Analysis of Subject Access Critiques". *Cataloging & Classification Quarterly* 32, no. 2: 61-80. https://doi.org/10.1300/J104v32n02_06

Pacey, Philip. 1989. "The Classification of Literature in the Dewey Decimal Classification: The Primacy of Language and the Taint of Colonialism". *Cataloging & Classification Quarterly* 9, no. 4: 101–7. https://doi.org/10.1300/J104v09n04_08

Price, Ludi. 2019. "Fandom, Folksonomies and Creativity: The Case of the Archive of Our Own". In *The Human Position in an Artificial World: Creativity, Ethics and AI in Knowledge Organization*, edited by David Haynes and Judi Vernau, 11–37. Germany: Ergon Verlag. https://doi.org/10.5771/9783956505508-11.

Schmidt, Kjeld and Liam Bannon. 1992. "Taking CSCW Seriously: Supporting Articulation Work". *Computer Supported Cooperative Work (CSCW)* 1: 7–40. https://doi.org/10.1007/BF00752449

Snow, Karen. 2017. "Defining, Assessing, and Rethinking Quality Cataloging". *Cataloging & Classification Quar-*

terly 55, no. 7–8: 438–55. https://doi.org/10.1080/01639374.2017.1350774.

Stahl, Lynne. 2024. "Uncontrollable Vocabularies: : Queer Theory, Sexual Identity, and the Catalog". *Journal of Radical Librarianship* 10: 17–35. https://journal.radicallibrarianship.org/index.php/journal/article/view/101

Star, Susan Leigh. 1999. "The Ethnography of Infrastructure." *American Behavioral Scientist* 43, no. 3: 377–91. https://doi.org/10.1177/00027649921955326

Star, Susan Leigh and Anselm Strauss. 1999. "Layers of Silence, Arenas of Voice: The Ecology of Visible and Invisible Work". *Computer Supported Cooperative Work (CSCW)* 8: 9–30. https://doi.org/10.1023/A:1008651105359

Watson, Brian M. 2020. "'There Was Sex but No Sexuality*:' Critical Cataloging and the Classification of Asexuality in LCSH". *Cataloging & Classification Quarterly* 58, no. 6: 547–65. https://doi.org/10.1080/01639374.2020.1796876

Knowledge Organization Systems Classifying Crimes of Violence Against Women, Homicide of Women and Feminicide: A Proposal

Rochelle Martins Alvorcem*, Gercina Ângela de Lima**, and Maria Cristina Vieira de Freitas***

* *** University of Coimbra- FLUC, Largo da Porta Férrea 3004-530 Coimbra, Portugal

** Federal University of Minas Gerais, Av. Pres. Antônio Carlos, 6627 – Pampulha, Belo Horizonte – MG, Brazil

*rochelle.alvorcem@gmail.com, **glima@eci.ufmg.br, ***mcvdefreitas@gmail.com

Rochelle Martins Alvorcem is a PhD student in Information Science at the University of Coimbra (PT), a master in Information Science at the Federal University of Santa Catarina (BR), and a Librarian graduated from the Federal University of Rio Grande do Sul (BR). She has worked since 2010 with information architecture and information and knowledge management at the Brazilian Agricultural Research Corporation. His current research interests include Knowledge Organization and Representation, Ontologies, Information Architecture, Digital Data Curation and Semantic Web.

Gercina Ângela de Lima is a full professor at the Information Science School of Federal University of Minas Gerais, Brazil. She holds a PhD in Information Science from UFMG(Brazil) and a Masters in Library Science and Services from Clark Atlanta University (USA). She is the coordinator of the Research Group MHTX Research Group – Conceptual Modeling for Organization and Representation of Hypertextual Information, and has a Research Scholarship, from The Brazilian Research Council (CNPq), since 2010. Her currents research interests include Organization and Representation of Knowledge, Hypertext, Digital Library and Web Semantic.

Maria Cristina Vieira de Freitas has a PhD and Masters in Information Science. She is an Assistant Professor at the Department of Philosophy, Communication and Information at the Faculty of Arts of the University of Coimbra, Director of the University Archive and the 25 de Abril Documentation Center. She is a researcher at the Center for Interdisciplinary Studies – CEIS20, member of the International Society for Knowledge Organization (ISKO) – Iberian Chapter, of the Group of Experts in Archives of the *Asociación de los Archiveros Latino-Americanos* (ALA) and of the International Council of Archives (CIA). She has more than a hundred published works and participation in evaluation committees, congresses, national and international master's and doctorate juries.

Alvorcem, Rochelle Martins, Gercina Ângela de Lima, and Maria Cristina Vieira de Freitas. 2024. "Knowledge Organization Systems Classifying Crimes of Violence Against Women, Homicide of Women and Feminicide: A Proposal". *Knowledge Organization* 51, no. 8: 667-685. 28 references. DOI:10.5771/0943-7444-2024-8-667.

Abstract: According to the World Health Organization (WHO), one in three women worldwide has experienced some form of physical or sexual violence. In Brazil, in 2022, about 30% suffered some form of gender-based violence, with the number of women's homicides (4,034) surpassing that of femicides (1,437). Misclassified crimes? In 2015, Law 13,104 amended article 121 of the Brazilian Penal Code, establishing femicide as a qualifying circumstance for the crime of homicide (when the act occurs against the woman because of her female sex condition). Having Information Science as a basis for promoting the development of methodologies and strategic solutions, conceptual and technical treatments on this phenomenon are necessary. Aiming to contribute guidance for typifying crimes of violence against women in filling out the National Risk Assessment Form (FNAR) and Police Report (BO) through Knowledge Organization Systems, it was based on Araújo and Lima (2018), who suggest semantic enrichment of ontologies through text corpus comparison, and applied METHODOE (Methodology for Domain Ontology Enrichment) by Araújo (2021), consisting of 5 steps. GSSO – Gender, Sex, and Sexual Orientation Ontology, Thesaurus for Gender and Women's Studies, Women Thesaurus, and Thesaurus on Violence against Women and Girls were used as a basis. The result was the OntoVDFcM ontology, to be used in filling out FNAR and BO forms, reducing errors, identifying or preventing crime more assertively, and contributing to

more precise statistics, enabling investment of more time and money in protective, preventive, supportive, educational actions, and mainly, in more effective public policies.

Received: 10 June 2024; Revised: 6 August 2024; Accepted: 21 August 2024.

Keywords: knowledge organization systems; feminicide; OntoVDFcM ontology.

1.0 Introduction

According to the World Health Organization (WHO), one in every three women worldwide (approximately 736 million) has experienced some type of physical or sexual violence. This occurs in all countries and cultures, causing irreparable damage to millions of women and their families (Nações Unidas 2021).

In Brazil, about 30% of Brazilian women have experienced some type of gender-based violence. The number of women's homicides was 4,034, while that of femicides was 1,437, meaning 180% more women's homicides (Bueno et al. 2023). Are we dealing with incorrectly classified crimes?

In 2015, Law 13,104 amended article 121 of the Brazilian Penal Code, establishing femicide as a qualifying circumstance for the crime of homicide when the act occurs against the woman because of her female sex condition (when the crime involves domestic violence, contempt, or discrimination against the woman's condition) (Brazil 2015).

At the time of reporting homicides of women and femicides, when filling out the occurrence records, both professionals and individuals complete some fields following the questionnaire flow, with fixed responses, and the only way to choose another response is: "None of the above aggressions," or "none of the above-listed behaviors," or "yes, otherwise". In other words, the lack of options and a methodology or standard complicates and distorts the content and analysis of the FNAR – National Risk Assessment Form, filled out when the death crime has not yet occurred, and the complaint in the BO – Police Report, filled out when the death crime has already occurred. It is perceived that we need semantic and terminological standards to be followed in filling out the forms, contemplating the domain "Domestic and family violence against women (VDFcM)".

The urgency and importance of the topic are evident as it is embedded in two of the 17 United Nations Sustainable Development Goals: SDG 5 – Achieving gender equality and empowering all women and girls; and SDG 16 – Promoting peaceful and inclusive societies for sustainable development, providing access to justice for all and building effective, responsible, and inclusive institutions at all levels (Nações Unidas 2022).

In this scenario, Information Science (Is) presents itself as support, interacting with Police Science (PS) at a technological level, promoting the development and selecting pertinent, useful, and valuable information for public safety, helping for the well-being of society (Moreira and Muriel-Torrado 2019).

By using knowledge organization systems (KOS), we can create a terminological control instrument with the purpose of aligning and standardizing the insertion of information in the complaint forms (FNAR and BO), promoting the development of solutions that support, through correct typification, the best information and interpretation of these situations.

Zeng (2008) explains that there are different types of KOS according to the complexity of their structures and functions, and adds that they can be described based on their structures (from planar to multidimensional), and their main functions. These functions include: eliminating ambiguity, controlling synonyms or equivalents, establishing explicit semantic relations, such as hierarchical and associative relationships, presenting both relationships and the properties of concepts in knowledge models. Examples of KOS include: term lists (lists, dictionaries, glossaries, synonym rings); similar models and metadata (authority files, directories); Classification and categorization (subject headings, taxonomies, categorization schemes, classification schemes); Relationship models (Thesauri, Semantic networks, and ontologies).

To elaborate a terminological control instrument to represent and share knowledge and be used in filling out the forms typified as FNAR and BO, in addition to contemplating the complexity of the theme and the different synonyms and ambiguities of the expressions presented in the Brazilian vocabulary, we need conceptual models in which the terms represent the concepts, their relationships, and synonyms. In this scenario, ontologies, as one of the models of KOS, present themselves as "specific conceptual models that represent complex relationships between objects, including rules and axioms missing in semantic networks" (Zeng 2008, 162).

According to Araújo (2021), domain ontologies are formatted from guidelines made by specialists, on the specific objectives of a knowledge domain, therefore, we will approach the use of domain ontologies and thesauri to conceptualize and classify terms from the perspective of Araújo and Lima (2018), Zeng (2008), and Campos (2010; 2017). In the complex purpose of elaborating consistent ontologies, definitions are fundamental, as they describe the semantic content of a term, providing possible semantic compatibility (Campos 2010). For good work in specific domains, it is important to use standards in definitory enunciation. Dahlberg (1983)

defines a definitive statement as a combination of attributes that describe a concept in a specific scenario. We will also use the onomasiological approach where the concept is the premise for the correct definition of terms and their relationships, Dahlberg's Concept Theory (1978), which emphasizes classification based on Aristotelian categories and Ranganathan's Faceted Classification Theory, according to the guidelines of Gomes et al. (2010) and Gomes and Campos (2019).

We sought in previous studies, research addressing the VDFcM theme in order to identify thesauri or ontologies that could represent such domain more realistically, being the most representative for this research: Thesaurus for Gender and Women's Studies (TPEDGESM) (Bruschini et al., 1998), with information in the field of gender studies, showing a "revolution" in terms of objectivity and reality about the female gender; Women's Thesaurus (Mediavilla Herreros and Folla Fernández 2014), elaborated from the collection of the Women's Library in Madrid, covering the evolution of the Feminist Movement in Spain, showing the contribution of women in the historical and social process, with a non-sexist language, aiming to make women visible, empowered, eliminating misogynistic terms in the designation of women; Thesaurus on Violence against Women and Girls (Morais, 2022), aiming at the democratization of utilitarian information for the protection, support, and support of women in situations of violence, contributing to the fight and prevention of gender violence; and finally, the Gender, Sex, and Sexual Orientation Ontology (GSSO) (EMBL-EBI, 2022), created in 2019, to provide a way to facilitate communication between LGBTQIA+ people and health professionals in the LGBTQIA+ domain, being classified among the top 5% of all ontologies in the BioPortal of the National Center for Biomedical Ontology (NCBO).

As this article is part of a research within the scope of an ongoing doctoral thesis, which studies the moment and the most appropriate way to typify homicides of women and femicides, we will focus on the study of KOS, more specifically on the enrichment of an ontology. Therefore, we will follow the methodology proposed by Araújo and Lima (2018), who suggest semantic enrichment of ontologies based on comparison to a text corpus, with the "objective of measuring the proximity between the ontology and the domain it proposes to represent" (649), complementing with Araújo (2021), the METHODOE (Methodology for Domain Ontology Enrichment).

2.0 Objectives

Our main goal is to provide guidance for better classification and prevention of women's homicides and feminicides through the development and application of SOC, with specific objectives including the semantic enrichment of the GSSO Ontology and the creation of the OntoVDFcM ontology, focused on Domestic and Family Violence against Women, based on the enriched GSSO Ontology.

3.0 Methodology and results

For the ontology to encompass the reality of the context under study, it is essential to pay special attention to its modeling and elaboration, thus using validated methodologies. According to Isotani and Bitencourt (2015), in addition to the conceptual analysis and modeling of the domain, identifying the central function of the ontology will be crucial, as well as reusing another ontology to support the development of the ontology of interest and evaluating, through competency questions, the correct representation of reality, concepts, and relationships.

In Figure 1, we present the methodological steps for developing an ontology through the semantic enrichment of another ontology. In this article, we reduce the presentation of the methodological steps to accommodate word limitations.

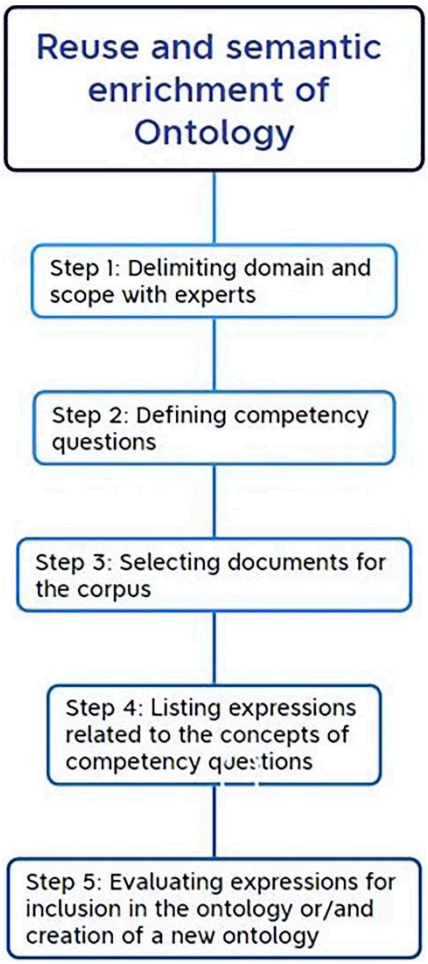


Figure 1. Reuse and semantic enrichment of Ontology – methodological steps.
Source: Prepared by the authors.

Given the proposal to enrich the GSSO ontology to more accurately encompass the domain of 'Domestic and Family Violence against Women', and following Campos et al. (2011), we have based our approach on domain ontologies. Domain ontologies aim to genuinely represent the characteristics of a conceptualization, portraying the proposed domain in a way that is closer to reality. These ontologies are developed based on the knowledge of domain experts, with well-defined purposes within a specific context. Furthermore, they provide greater semantic richness by constraining interpretations of concepts through a vocabulary and offer an ontological structure capable of creating a corpus suitable for validating the established conceptual model. To begin enriching the GSSO ontology following part of the methodology proposed by Araújo (2021), METHODOE, we will have five steps shown in Figure 1 and presented below: step 1– Delimiting domain and scope with experts; step 2 – defining competence issues; step 3 – selecting documents for the corpus; step 4 – listing expressions related to the concepts of competence issues; step 5 – evaluating expressions for inclusion in the ontology or/and creation of a new ontology.

3.1 Delimiting domain and scope with experts

It is important to answer some questions related to domain and scope, with the following steps:

a) Identification of the problem with specialists: classification of homicides of women and feminicides;
b) Definition of the ontology's theme: Domestic and Family Violence against Women – VDFcM;
c) Application and use of the ontology: within the context of Brazilian Civil Police in FNAR and BO forms;
d) Ontology for comparison and reuse: GSSO – Gender, Sex, and Sexual Orientation Ontology.

The crime of femicide is the end of a stage of various types of announced violence, which culminates in the death of a woman. After interviewing civil police officers who are experts in assisting women in situations of violence, eight CQs were formulated based on the research problem, the Brazilian Penal Code, and the five forms of domestic and family violence according to Article 7 of Law No. 11.340/2006 – Maria da Penha Law (Brazil 2006):

a) What is understood as domestic and family violence against women?
b) What is understood as femicide?
c) What is understood as the homicide of a woman?
d) What is understood as physical violence, what other terms can we attribute to this type of violence, and what examples can we consider?
e) What is understood as psychological violence, what other terms can we attribute to this type of violence, and what examples can we consider?
f) What is understood as economic violence, what other terms can we attribute to this type of violence, and what examples can we consider?
g) What is understood as moral violence, what other terms can we attribute to this type of violence, and what examples can we consider?
h) What is understood as sexual violence, what other terms can we attribute to this type of violence, and what examples can we consider?

In the verification of the eight competency questions, 11 terms were identified that will guide the answers: domestic and family violence against women, domestic violence, family violence, violence against women, femicide, homicide of a woman, physical violence, psychological violence, economic violence, moral violence, and sexual violence.

3.2 Selecting documents for the corpus

When using documents such as laws, the Brazilian Penal Code, and national guidelines (Table 1), we will make ontology enrichment more effective and bring us closer to reality. According to Campos (2017), by using a consistent terminological corpus that accurately describes the domain, we can produce more complete ontologies.

The selected documents are those that support the legal decisions and actions of the Public Ministry and the Police, chosen for this stage of the research. As the research is ongoing, it is expected, in the future, to add other laws to compose a more complete picture of reality.

3.3 Selecting expressions related to concepts

The information extraction technique was manual, given that our Textual Corpus is made up of legally validated documents such as laws, decrees, standards, etc., and thus we can carry out conceptual and relationship enrichment (Araújo 2021). We present in Table 2, the terms chosen in the QCS, the number of definitory texts related to each of the terms chosen in the documents.

The definition increases the semantic expressiveness of ontologies, enabling semantic harmonization, as it describes the semantic content of a term. "This description allows intelligent agents to understand the meaning of a term and establish inferences about these meanings, as the definition is composed of characteristics of concepts, which are also concepts that are related to form the semantic understanding of the terms in question" (Campos 2010, 222).

Dahlberg (1978, 106) explains that "definitions are indispensable assumptions in argumentation and verbal commu-

Law	Date	Details
Decree Law nº 2.848	07/12/1940	Brazilian Penal Code
Decree nº 19.841	22/10/1945	Promulgates the United Nations Charter, of which the Statute of the International Court of Justice, signed in San Francisco on June 26, 1945, is an integral part, on the occasion of the Conference of the International Organization of the United Nations.
Universal Declaration of Human Rights	1948	The document (Resolution 217 A III) was adopted by the United Nations General Assembly in 1948, during a period of restructuring of the post-Second World War world order, and served as the basis for the drafting of several constitutional charters around the world, including the current Brazilian Federal Constitution.
Decree nº 52.476	12/09/1963	Promulgates the Convention on the Political Rights of Women, adopted on the occasion of the Seventh Session of the United Nations General Assembly.
Decree nº 89.460	20/03/1984	Promulgated the Convention on the Elimination of All Forms of Discrimination Against Women/CEDAW, 1979
Inter-American Commission on Human Rights	1994	Inter-American Convention to Prevent, Punish, and Eradicate Violence against Women, "Convention of Belém do Pará"
Decree nº 1.973	01/08/1996	Decree nº 1,973, of 01/08/1996, which promulgated the Inter-American Convention on Preventing, Punishing, and Eradicating Violence against Women, Belém do Pará, 06/09/1994
Law nº 11.340	07/08/2006	Creates mechanisms to combat domestic and family violence against women, in accordance with § 8 of art. 226 of the Federal Constitution, the Convention on the Elimination of All Forms of Discrimination against Women, and the Inter-American Convention to Prevent, Punish, and Eradicate Violence against Women; provides for the creation of Special Courts for Domestic and Family Violence against Women; amends the Code of Criminal Procedure, the Penal Code, and the Penal Execution Law; and provides other measures. Maria da Penha Law.
Law nº 13.104	09/03/2015	Amends art. 121 of Decree-Law nº 2,848, of December 7, 1940 – Penal Code, to provide for femicide as a qualifying circumstance of the crime of homicide, and art. 1 of Law nº 8,072, of July 25, 1990, to include femicide in the list of heinous crimes. Femicide Law
Women's Legislation	2016 (7th ed.)	Chamber of Deputies (Legislation Series, 230)
UN Women	2016	National Guidelines for Femicide: investigating, prosecuting, and judging violent deaths of women

Table 1. Laws on the topic of Domestic and Family Violence against Women.
Source: prepared by the authors based on the aforementioned legislation.

nications and that they constitute necessary elements in the construction of scientific systems". The author states that the characteristics, in addition to presenting the relationships between concepts, determine the definitions, and are divided into essential (necessary and sufficient), and accidental (optional) (Dahlberg 1983).

Araújo (2021) infers about the importance of concepts in the conception of SOC, because based on the principles of Dahlberg's Concept Theory (1978), when we apply consistent definitions in ontologies, we will be able to establish relationships (relationships or properties) between classes and subclasses, making -the most specific and assertive ones, and in the elaboration of axioms (restrictions/rules on relationships, allowing the use of inferences that do not appear in the Thesauruses).

When selecting candidate knowledge for the ontology, resulting from the application of the extraction technique, definitory enunciation were sought and relevant concepts were selected that will be inserted in the context of OntoVDFcM (Table 3).

The definition is a limitation of the idea of the concept, and only general concepts need to be defined, as they imply differences from other concepts. Individual concepts, on the other hand, have their own well-established objects (they have the form of time and space). But when we identify approximately 200 relevant concepts, referring to the general concepts extracted from the QCs, it is important to carry out an analysis so that we can extract all the necessary characteristics and help compose the general concepts (Dahlberg 1978).

3.4 Evaluation of expressions and inclusion in the ontology

For ontologies to represent knowledge, it is necessary to identify their general and auxiliary components. The gen-

QC No.	Search term/expression	Amount recovered	Definitory enunciation	Located in documents
QC-1	domestic and family violence counts women	34	1 (Law No. 11,340)	Decree Law No. 2,848; Law No. 11,340
QC-1	violence against women	35	1 (Decree nº 1,973)	Decree Law nº 2,848; Decree No. 1,973, Law No. 11,340
QC-1	Domestic violence	70	1 (Decree Law No. 2,848)	Decree Law No. 2,848; Law No. 11,340
QC-2	Feminicide	3	1(Decree Law nº 2,848/Law nº 13,104)	Decree Law nº 2,848; Law No. 13104
QC-3	Homicide of a woman	0	1(Decree Law nº 2,848/Law nº 13,104)	Decree Law No. 2,848, of December 7, 1940 – Brazilian Penal Code, Section VI – added by Law No. 13,104, of 3/9/2015
QC-4	physical violence	2	1(Law nº 11,340)	Law nº 11,340
QC-5	psychological violence	2	2 (Decree Law nº 2,848; Law nº 11,340)	Decree Law nº 2,848; Law No. 11,340
QC-6	property violence	1	1 (Law nº 11,340)	Law nº 11,340
QC-7	moral violence	1	1 (Law nº 11,340)	Law nº 11,340
QC-8	Sexual violence	2	1 (Law nº 11,340)	Law nº 11,340

Table 2. Definitory enunciation recovered in the collection of legal documents consulted.
Source: prepared by the authors based on the aforementioned legislation.

eral components: Classes and subclasses (elements composed of similar attributes represent the concepts of the domain, and the basic units of an ontology); properties/attributes (characteristics/qualities of classes and subclasses); relations (relationship or property, responsible for the semantic links between concepts in a given domain); axioms or restrictions (rules about relationships, allowing inferences to be made and restricting interpretations so that there are no ambiguities; bring the ontology closer to the intended model; and instances (or individuals) (represent a given object of a concept) (Araújo 2021).

For each of the expressions originating from the QCs, information was sought in the GSSO (1), TPEDGESM (2), Tesauro de Mujeres (3), and Tesauro sobre Violência contra Mulheres e Meninas (4), to identify existing terms, their relationships and definitions. Following Araújo (2021), we created a table showing concepts (conceptual enrichment), synonymous terms (lexical enrichment), relationships between concepts (relationship enrichment), resulting in, in addition to the table below, a glossary of selected terms (Appendix A). We will only show two terms, pointing with numbers, when in any of the SOCs, signaling those that need to be enriched in the definitions and consequently in the relationships (Table 4). As the ontology is expected to be used by Brazilian civil police, the use of the Portuguese language (BR) is default, and it is important to add terms in Portuguese to the VDFcM enriched ontology.

To ensure that the information complies with the objective of the OntoVDFcM ontology and the purpose of enrichment, the list will be presented to experts for validation and changes (if necessary). In case of changes, the list will be presented again after modifications.

3.5 Creation of the OntoVDFcM ontology

To add nee classes, subclasses and properties, we present a sample concept map (Figure 2) of the enriched OntoVDFcM ontology. We will only present the map with the term Patrimonial Violence, which I not included in the GSSO.

Figure 3 shows a sample of intermediate representations (relations), presented in triplets, in natural language (LN).

To formalize and implement the OntoVDFcM ontology, based on the conceptual model represented in Figure 3 in LN, we will transform knowledge into formal language using computer-readable language, OWL (Ontology Web Language), through the Protége software, an ontology editor. Following the logic proposed by Araújo (2021), we will do it in the following order: insertion of concepts such as classes and subclasses; definitions and synonyms of classes; creation of relationships and their logical properties; and the axioms between classes. We present the ontology in Figure 4.

At the end of the thesis, we will make available a guiding document with steps and modifications made so that we can share it between the Brazilian Civil and Military Police.

QC No	Search term/ expression	Definitory enunciation	Relevant concepts
QC-1	Domestic and Family Violence Against Women	Article 5. For the purposes of this Law, domestic and family violence against women is defined as any action or omission based on gender that causes her death, injury, physical, sexual, or psychological suffering, as well as moral or patrimonial harm: (See Supplementary Law No. 150, 2015); I – within the domestic unit, understood as the permanent living space of individuals, with or without family ties, including those sporadically aggregated; II – within the family, understood as the community formed by individuals who are or consider themselves related, united by natural, affinity, or express will bonds; III – in any intimate relationship of affection, in which the aggressor lives or has lived with the victim, regardless of cohabitation. Sole paragraph. The personal relationships referred to in this article are independent of sexual orientation. Article 6. Domestic and family violence against women constitutes one of the forms of violation of human rights.	Action, omission, gender, death, injury, physical suffering, sexual suffering, psychological suffering, moral harm, patrimonial harm, domestic unit, permanent coexistence, individuals, family ties, sporadically aggregated, family, community, related individuals, united, natural bonds, affinity, express will, intimate relationship, relationship of affection, aggressor, victim, cohabitation, personal relationships, sexual orientation, violation of human rights
QC-1	Violence Against Women	Violence against women is any action or conduct, based on gender, that causes death, harm, or physical, sexual, or psychological suffering to women, both in the public and private spheres. It is understood that violence against women includes physical, sexual, and psychological violence: §1. That occurred within the family or domestic unit or in any other interpersonal relationship, in which the aggressor lives or has lived in the same household as the woman and which includes, among others, rape, violation, mistreatment, and sexual abuse: §2. That occurred in the community and is perpetrated by any person and which includes, among others, rape, sexual abuse, torture, mistreatment of people, trafficking of women, forced prostitution, kidnapping, and sexual harassment in the workplace, as well as in educational institutions, health establishments, or any other place, and §3. That is perpetrated or tolerated by the State or its agents, wherever it occurs.	Physical violence, sexual violence, psychological violence, patrimonial violence, moral violence, action, conduct, gender, death, harm, physical suffering, sexual suffering, psychological suffering, public sphere, private sphere, within the family, domestic unit, interpersonal relationship, aggressor, household, woman, rape, violation, mistreatment, sexual abuse; community, people, torture, trafficking of women, forced prostitution, kidnapping, sexual harassment, workplace, educational institutions, health establishments, state, state agents
QC-1	Domestic Violence	§ 9. If the injury is committed against an ascendant, descendant, sibling, spouse or partner, or with whom they live or have lived, or, still, taking advantage of domestic relations, cohabitation, or hospitality relationships: Penalty – detention, from 3 (three) months to 3 (three) years. (Paragraph added by Law No. 10,886, of 7/17/2004, and with wording given by Law No. 11,340, of 8/7/2006, published in the DOU of 8/8/2006, in force 45 days after publication) § 10. In the cases provided for in paragraphs 1 to 3 of this article, if the circumstances are those indicated in § 9 of this article, the penalty is increased by 1/3 (one third). (Paragraph added by Law No. 10,886, of 7/17/2004) § 11. In the hypothesis of § 9 of this article, the penalty shall be increased by one third if the crime is committed against a disabled person. (Paragraph added by Law No. 11,340, of 8/7/2006, published in the DOU of 8/8/2006, in force 45 days after publication) § 12. If the injury is committed against an authority or agent described in arts. 142 and 144 of the Federal Constitution, members of the prison system and	Injury, ascendant, descendant, sibling, spouse, partner, live together, agent, domestic relations, cohabitation, hospitality, detention, penalty, crime, disabled person, injury, authority, agent, members of the prison system, members of the National Public Security Force, exercise of function, against spouse, as a result of the function, woman, injury committed against a woman, female sex condition, female sex.

Table 3. Definitory Enunciation and relevant concepts *(to be continued).*
Source: ONU Mulheres (2016, 125-127), National risk assessment form – domestic and family violence against women (Conselho Nacional de Justiça 2020), *Decreto-lei no 2.848, de 7 de dezembro de 1940* (Brazil 1940), Maria da Penha Law (Brazil 2006).

QC No	Search term/ expression	Definitory enunciation	Relevant concepts
		the National Public Security Force, in the exercise of their function or as a result of it, or against their spouse, partner or blood relative up to the third degree, by virtue of this condition, the penalty is increased by one to two thirds. (Paragraph added by Law No. 13,142, of 7/6/2015) § 13. If the injury is committed against a woman, for reasons of female sex, as provided for in § 2-A of art. 121 of this Code: Penalty – imprisonment, from 1 (one) to 4 (four) years. (Paragraph added by Law No. 14,188, of 7/28/2021).	
QC-2	Femicide	(Legal name added by Law No. 13,104, of 3/9/2015) Qualified Homicide VI – against women for reasons of female sex condition; (Paragraph added by Law No. 13,104, of 3/9/2015) § 2. - A, there are reasons of female sex condition when the crime involves: I – domestic and family violence; II – disregard or discrimination against the female condition.	Qualified homicide, woman, against women, female sex, domestic and family violence, disregard, discrimination, female condition, disregard for the female condition
QC-3	Woman Homicide	Killing someone – qualified homicide: if the homicide is committed: I – for pay or promise of reward, or for another vile reason; II – for futile reason; III – with the use of poison, fire, explosive, suffocation, torture or other insidious or cruel means, or that may result in common danger; IV – treacherously, by ambush, or by dissimulation or other resource that hinders or makes impossible the defense of the victim; V – to ensure the execution, concealment, impunity or advantage of another crime;	Qualified homicide, woman, against women, female sex, domestic and family violence, disregard, discrimination, female condition, disregard for the female condition
QC-4	Physical Violence	Article 7. Forms of domestic and family violence against women include, among others: I – physical violence, understood as any conduct that offends her integrity or physical health; Types of physical violence: Drownings; Hand assaults; Object assaults; Threats to children; Death threats; Suicide threats; Threats of physical violence; Throwing objects; Lack of communication as a form of punishment: does not listen or speak; Pinches; Kicks; Decides for her; Destruction of objects with special sentimental value Pushes; Beatings; Strangulation; Injuries caused by burns; Injuries caused by firearms; Injuries caused by knives or cutting objects; Shaking and squeezing arms; Prevents or hinders access to work, study or any other activity; Injuries with cutting objects; Injuries with piercing objects; Mistreatment of pets; Bites; Sudden and unjustified mood changes, in the face of the same situation or behavior: both praise and humiliate; Does not allow the victim to make decisions about her life or family, nor participate in decisions; Beatings; Deprivation of basic needs (food, sleep, etc.); Hair pulls; Guilt and confusion, caused by constant corrections and manifestations made by the abuser, with the intention of presenting themselves as victims; Punches; Suffocation; Hanging; Slaps; Torture.	Conduct, offend, bodily integrity, physical health, Drownings, assaults, Hand assaults; Object assaults; Threats, Threats to children; Death threats; Suicide threats; Threats of physical violence; Throwing objects; Lack of communication, punishment, does not listen, does not speak, Lack of communication as a form of punishment: does not listen or does not speak; Pinches; Kicks; Decides for her; Destruction of objects with special sentimental value, destruction, objects of special value, objects of value, Pushes; Beatings; Strangulation; Injuries caused by burns; Injuries caused by firearms; Injuries caused by knives or cutting objects; Shaking and squeezing arms; Prevents access to work; hinders access to work; Prevents or hinders access to work, study or any other activity; prevents access to study, access to study, injuries, cutting objects; Injuries with cutting objects; piercing

Table 3 continued

QC No	Search term/ expression	Definitory enunciation	Relevant concepts
			objects; Mistreatment of pets; Bites; Sudden and unjustified mood changes, in the face of the same situation or behavior: sudden changes, both praise and humiliate; praise, humiliations; Does not allow the victim to make decisions about her life or family, nor participate in decisions; Beatings; Deprivation of basic needs (food, sleep, etc.); Deprivations, basic needs, food, sleep, Hair pulls; Guilt and confusion, caused by constant corrections and manifestations made by the abuser, with the intention of presenting themselves as victims: Punches; Suffocation; Hanging; Slaps; Torture.
QC-5	Psychological Violence	Article 7. Forms of domestic and family violence against women include, among others: II – psychological violence, understood as any conduct that causes emotional harm and decreases self-esteem or that harms and disturbs the full development or that aims to degrade or control her actions, behaviors, beliefs and decisions, through threats, coercion, humiliation, manipulation, isolation, constant surveillance, chronic persecution, insult, blackmail, violation of her privacy, ridicule, exploitation and limitation of the right to come and go or any other means that cause harm to her psychological health and self-determination; (Text given by Law No. 13,772, 2018). Types of psychological violence: Alienation; Abandonment; Threats; Intimidating physical approach; Knocking on doors; Blackmail; Embarrassment; distorting and omitting facts to leave the woman in doubt about her memory and sanity (gaslighting); examining her drawers and belongings. Exploitation; Screaming; Humiliation; Insistence on considering the victim crazy, stupid, or useless; Insults; Isolation; Limitation of the right to come and go or any other means that cause harm to her psychological health and self-determination; Insults; Manifesting jealousy and continuous suspicions; Manipulation; Persecution; Prohibiting studying; Prohibiting traveling; Prohibiting from talking to friends; Prohibiting from talking to relatives; Prohibiting from leaving home; Ridicule; Restricting freedom of belief; Constant surveillance.	Conduct, emotional harm, Alienation; Abandonment; Threats; Intimidating physical approach; Knocking on doors; Blackmail; Embarrassment; Distorting and omitting facts to leave the woman in doubt about her memory and sanity (gaslighting); Examining her drawers and belongings; Exploitation; Screaming; Humiliation; Insistence on considering the victim crazy, stupid, or useless; Insults; Isolation; Limitation of the right to come and go or any other means that cause harm to her psychological health and self-determination; Insults; Manifesting jealousy and continuous suspicions; Manipulation; Persecution; Prohibiting studying; Prohibiting traveling; Prohibiting from talking to friends; Prohibiting from talking to relatives; Prohibiting from leaving home; Ridicule; Restricting freedom of belief; Constant surveillance.
QC-6	Patrimonial Violence	Article 7. Forms of domestic and family violence against women include, among others: IV – patrimonial violence, understood as any conduct that constitutes retention, subtraction, partial or total destruction of her objects, work instruments, personal documents, assets, values and rights or economic resources, including those intended to meet her needs; Types of patrimonial violence: Causing intentional damage to objects belonging to the woman or those she cares about; Controlling money; Control over work and salary (it can even	Causing intentional damage to objects belonging to the woman or those she cares about; Controlling money; Control over work and salary (it can even involve withholding money); Failure to pay alimony; Destruction of personal documents; Fraud; Theft, extortion or damage; Not providing access to the money

Table 3 continued

QC No	Search term/ expression	Definitory enunciation	Relevant concepts
		involve withholding money); Failure to pay alimony; Destruction of personal documents; Fraud; Theft, extortion or damage; Not providing access to the money necessary to meet the family's needs; Deprivation of goods, values, or economic resources; Subtraction or destruction of belongings, objects, personal documents, assets, values, resources, etc., belonging to the woman, to cause her harm and control her.	necessary to meet the family's needs; Deprivation of goods, values, or economic resources; Subtraction or destruction of belongings, objects, personal documents, assets, values, resources, etc., belonging to the woman, to cause her harm and control her;
QC-7	Moral Violence	Article 7. Forms of domestic and family violence against women include, among others: V – moral violence, understood as any conduct that constitutes slander, defamation, or insult. Types of moral violence: Accusing the woman of betrayal; reporting the victim to the police; devaluing the victim for her way of dressing; distorting and omitting facts to leave the woman in doubt about her memory and sanity; passing moral judgments on behavior; exposing intimate life; making false criticisms; Playing the victim in public, claiming that she mistreats him; preventing or hindering external relationships to the couple; preventing her from accompanying him in activities or imposing his presence by force; Social isolation; belittling the woman through insults that target her character; trying to tarnish the woman's reputation	Accusing the woman of betrayal; Reporting the victim to the police; Devaluing the victim for her way of dressing; Distorting and omitting facts to leave the woman in doubt about her memory and sanity; Passing moral judgments on behavior; Exposing intimate life; Making false criticisms; Playing the victim in public, claiming that she mistreats him; Preventing or hindering external relationships to the couple; Preventing her from accompanying him in activities or imposing his presence by force; Social isolation; Belittling the woman through insults that target her character; Trying to tarnish the woman's reputation
QC-8	Sexual Violence	Article 7. Forms of domestic and family violence against women include, among others: III – sexual violence, understood as any conduct that constrains her to witness, maintain, or participate in unwanted sexual relations, through intimidation, threat, coercion, or the use of force; that induces her to commercialize or use her sexuality in any way, that prevents her from using any contraceptive method, or that forces her into marriage, pregnancy, abortion, or prostitution, through coercion, blackmail, bribery, or manipulation; or that limits or annuls the exercise of her sexual and reproductive rights; Types of sexual violence: Coercion to maintain sexual relations, using physical force or emotional blackmail; Rape; Forcing marriage, pregnancy, or prostitution through coercion, blackmail, bribery, or manipulation; Humiliations related to her sexual conduct; Preventing the use of contraceptives or forcing the woman to have an abortion; Limiting or nullifying the exercise of the woman's sexual and reproductive rights; Forcing the woman to commercialize or use her sexuality in any way; Forcing the woman to perform sexual acts that cause discomfort or repulsion; Degrading treatment of females; Violence and assaults during pregnancy.	Intimidation, threat, use of force, Coercion to maintain sexual relations, using physical force or emotional blackmail; Rape; Forcing marriage, pregnancy, or prostitution through coercion, blackmail, bribery, or manipulation; Humiliations related to her sexual conduct; Preventing the use of contraceptives or forcing the woman to have an abortion; Limiting or nullifying the exercise of the woman's sexual and reproductive rights; Forcing the woman to commercialize or use her sexuality in any way; Forcing the woman to perform sexual acts that cause discomfort or repulsion; Degrading treatment of females; Violence and assaults during pregnancy.

Table 3 continued

QC No.	Concept (conceptual enrichment)	Thesauri & GSSO	Synonymous terms (lexical enrichment)	Relationships with other concepts (relationship enrichment)
QC-1	Domestic and family violence Against the woman	1, 2, 3	Domestic violence against women; Family violence against women; Abuse against women, gender-based violence; marital violence; violence in the home; intrafamily violence; domestic violence; relationship violence; violence within the family, violence in the domestic environment; violence in marriage; Dating violence/domestic abuse; mistreatment of women; domestic assault against women; family violence	isPracticedBy: people with family ties, people without family ties, partners isType: violence, aggression, physical suffering. Sexual suffering isBasedOn: gender isCausedBy: gender-based action or inaction doesnotDependOn sexual orientation resultsIn: death, injury, physical suffering, sexual suffering, psychological suffering, moral damage, property damage hasAsContext: scope of the family, family environment, scope of the domestic unit, intimate relationship of affection constitutesAFormOf: violation of human rights hasImpact: Physical, psychological, social impact affectsThe: family, the wife, the children isAProductOf: neglect, aggression, abuse, violence, mistreatment, discrimination, exploitation, cruelty, oppression isATypeRelatedTo: Violence is a form of violence, physical violence, domestic violence hasLaws: Law No. 11,340, 2006 hasInjuries: bruises, cuts, bruises hasLegalConsequences: suspension of possession or restriction of carrying of weapons, removal from the home, domicile or place of coexistence with the offended party, prohibition of approaching the offended party, her family and witnesses hasMinimumDistanceBetweenAggressorAndVictim: to be fixed hasHowTypes: physical violence, psychological violence, moral violence, sexual violence, property violence hasInstitutionThatPrevents: Union, State, Federal District, Municipalities, Judiciary, Public Ministry, Public Defender hasAreasInInstitutionsForPrevention: Public security, social assistance, health, education, work and housing, women's assistance police stations, hasPreventiveMeasure: Public policies, non-governmental actions

Table 4. Organization of acquired knowledge *(to be continued)*.
Source: ONU Mulheres (2016, 125-127), National risk assessment form – domestic and family violence against women (Conselho Nacional de Justiça 2020), *Decreto-lei no 2.848, de 7 de dezembro de 1940* (Brazil 1940), Maria da Penha Law (Brazil 2006).

QC No.	Concept (conceptual enrichment)	Thesauri & GSSO	Synonymous terms (lexical enrichment)	Relationships with other concepts (relationship enrichment)
				hasProtectionMeasure: urgent protective measures hasSynonyms: Domestic violence against women; Family violence against women; Abuse against women, gender-based violence; marital violence; violence in the home; intrafamily violence; domestic violence; relationship violence; violence within the family, violence in the domestic environment; violence in marriage; Dating violence/domestic abuse; mistreatment of women; domestic assault against women; family violence
QC-6	property violence	4	Property damage, Financial violence, Financial abuse, Control of assets, Destruction of assets, Destruction of personal property, retention of assets, procedural plunder	isPracticedBy: people with family ties, people without family ties, partners isTypeof: Domestic and family violence against women, Violence against women, financial violence, property damage, financial abuse, Control of assets, Destruction of assets, Destruction of personal property, retention of property, procedural embezzlement isBasedOn: gender isCausedBy: gender-based action or inaction doesnotDependOn: sexual orientation resultsIn: death, injury, property damage hasAsContext: scope of the family, family environment, scope of the domestic unit, intimate relationship of affection constitutesAFormOf: violation of human rights hasImpact: Physical, psychological, social affectsThe: family, the wife, the children isAProductOf: neglect, aggression, abuse, violence, mistreatment, discrimination, exploitation, cruelty, oppression isTypeRelatedTo: Violence isAFormOf: violence, physical violence, domestic violence hasLaw: Law No. 11,340, 2006 hasLegalConsequences: suspension of possession or restriction of carrying of weapons, removal from the home, domicile or place of coexistence with the offended party, prohibition of approaching the offended party, her family and witnesses hasMinimumDistanceBetweenAggressorAndVictim: to befixed

Table 4 continued

QC No.	Concept (conceptual enrichment)	Thesauri & GSSO	Synonymous terms (lexical enrichment)	Relationships with other concepts (relationship enrichment)
				hasLegalDefinition: Any conduct that constitutes retention, subtraction, partial or total destruction of your objects, work instruments, personal documentation, goods, values and rights or economic resources, including those intended to satisfy your needs; act of violence that involves damage, loss, subtraction, destruction or retention of objects, personal documents, goods and valuables. hasSynonym: Emotional and Patrimonial Violence hasHowTypes: Purposefully causing damage to objects belonging to the woman or things she likes, Controlling money, Control over work and salary (may even withhold money), Failure to pay alimony, Destruction of personal documents, Embezzlement, Theft, extortion or damage, Not providing access to the money necessary to meet the family's needs, Depriving of goods, values or economic resources, Subtraction or destruction of goods, objects, personal documents, assets, values, resources, etc., belonging to the woman, to cause harm and control it

Table 4 continued

The choice of the GSSO ontology allowed us to reach 70% of the initial terms that will be available in OntoVDFcM. With regard to QCs, two important classes (property violence and moral violence) were not included in the GSSO. The vast majority of terms found present definitions, but need complementation, as most definitions for the VDFcM context are based on Laws and decrees. Regarding relationships such as object properties and data properties, it was found that they do not exist in GSSO and were added in OntoVDFcM.

We present the first results, as the subject is very dense and the samples are large. Another important point is that the GSSO ontology is entirely in English. In the case of the new OntoVDFcM ontology, which we intend to use in Brazil, it results in terms translated and defined in the Portuguese language – BR.

5.0 Conclusions

IC is a human and social science, in which information is treated as "something alive", dynamic, relating to society, at a specific moment, and in a given context (Araújo and Lima 2018). In the same way, ontologies are also, as they need to be updated, fed and adjusted to the context. When using domain ontologies, we thought about the potential for representing knowledge in a reliable way with regard to the intended domain, being readable by both humans and machines (using languages such as OWL/RDF).

In Table 1, documents composed of laws and resolutions that deal with Women's Rights were presented. In Table 2, the number of definitory Enunciation found in the documents was presented. In Table 3, the relevant concepts arising from the definitory Enunciation were presented, which were compared to the GSSO ontology and thesauruses. Such concepts need to be aligned according to the VDFcM context. In Table 4, a small sample of the organization of knowledge acquired from QCs was listed. We used 3 types of enrichment proposed by Araújo (2021): conceptual, lexical and relations. From this table, a conceptual map was created (Figure 2), showing the proposed relationships. In Figure 3, the elaborated relationships were shown in LN, ready to be added to the new ontology. And finally, a preview of

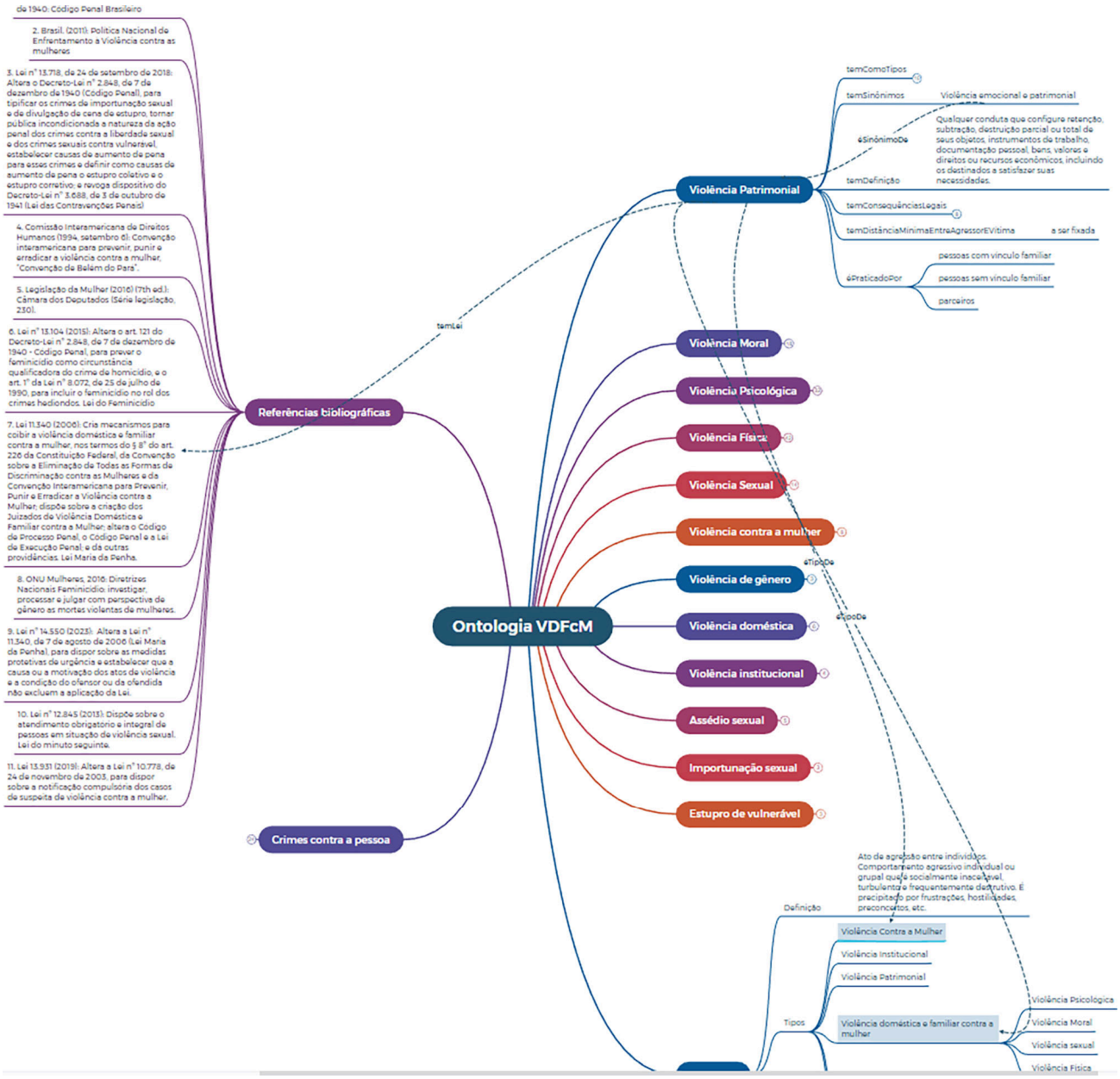


Figure 2. Conceptual Map of the GSSO + OntoVDFcM Ontology – Patrimonial Violence.
Source: prepared by the authors

the new ontology in the Protége software was shown. (OntoVDFcM.rdf) OWX

According to Meneghel and Portella (2017), the crime of femicide is the final stage in an escalation of violence against women, knowing that the vast majority of these deaths are "announced" and could be avoided.

This difference between records of female homicides and femicides demonstrates how public policy does not depend solely on the approval of a law. It is necessary to record and monitor data, in addition to care in implementation, which, in this case, seems to be lacking.

It is expected that the use of the OntoVDFcM ontology, when filling out the occurrence (FNAR or BO), will help in preventing a possible crime, or at least in identifying the crime correctly. Using the ontology, when filling in the fields, the record options will be listed standardized and within the VDFcM knowledge domain. In this way, we will contribute so that statistics portray the real situation and act in favor of the urgent need to invest more time and money in protection, prevention, reception, education, and public policies. And above all, it is expected that the Civil Police realizes the importance of this work and use it.

Term	relationship	Term Y
Violence	HasHowTypes	Violence against women
Violence against women	HasHowTypes	Moral Violence
Violence against women	HasHowTypes	Sexual Violence
Violence against women	HasHowTypes	Psychological violence
Violence against women	HasHowTypes	Physical violence
Violence against women	HasHowTypes	Patrimonial Violence
Patrimonial Violence	affectsA	family
Patrimonial Violence	affectsA	the woman
Patrimonial Violence	affectsA	the children
Patrimonial Violence	constitutesAFormOf	violation of human rights
Patrimonial Violence	isBasedOn	gender
Patrimonial Violence	isCausedBy	gender-based omission
Patrimonial Violence	isCausedBy	gender-based action
Patrimonial Violence	isPracticedBy	people with family ties
Patrimonial Violence	isPracticedBy	people without family ties
Patrimonial Violence	isPracticedBy	partners
Patrimonial Violence	isProductOf	negligence
Patrimonial Violence	isProductOf	aggression
Patrimonial Violence	isProductOf	abuse
Patrimonial Violence	isProductOf	violence
Patrimonial Violence	isProductOf	mistreatment
Patrimonial Violence	isProductOf	discrimination
Patrimonial Violence	isProductOf	exploration
Patrimonial Violence	isProductOf	cruelty
Patrimonial Violence	isProductOf	oppression
Patrimonial Violence	isType	Violence against women
Patrimonial Violence	isType	Domestic and family violence against women
Patrimonial Violence	isType	financial violence
Patrimonial Violence	isType	property damage
Patrimonial Violence	isType	financial abuse
Patrimonial Violence	isType	Asset control
Patrimonial Violence	isType	Destruction of property
Patrimonial Violence	isType	Destruction of personal property
Patrimonial Violence	isType	asset retention
Patrimonial Violence	isType	waste procedural
Patrimonial Violence	isTypeRelatedA	violence
Patrimonial Violence	it's a way of	violence
Patrimonial Violence	it's a way of	Physical violence
Patrimonial Violence	it's a way of	domestic violence

Figure 3. OntoVDFcM relationships.
Source: prepared by the authors

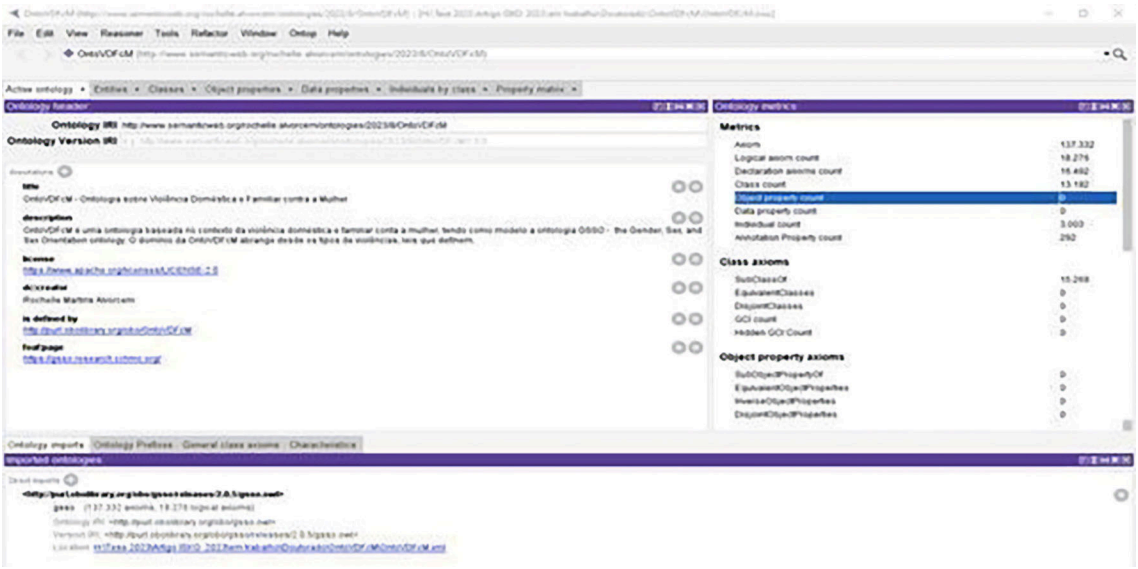


Figure 4. OntoVDFcM Ontology.

References

Araújo, W. J. 2016. "Avaliação de ontologia com base na comparação a um corpus: um estudo da OntoAgroHidro da EMBRAPA". Master's Thesis, Universidade Federal de Minas Gerais.

Araújo, W. J. 2021. "Proposta metodológica para enriquecimento de ontologias de domínio". PhD diss., Universidade Federal de Minas Gerais.

Araújo, Webert Júnio and Gercina Ângela de Lima. 2018. "Avaliação e enriquecimento semântico de ontologia com base na comparação a um corpus: um estudo de caso no domínio da agricultura". *Revista Ibero-Americana de Ciência da Informação* 11, no. 3: 645–62. https://doi.org/10.26512/rici.v11.n3.2018.10433

Brazil. 1940. *Decreto-lei no 2.848, de 7 de dezembro de 1940*. http://www.planalto.gov.br/ccivil_03/Decreto-Lei/Del2848.htm.

Brazil. 2015. *Lei no 13.104, de 9 de março de 2015*. http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2015/Lei/L13104.htm#art1.

Brazil. 2006. *Lei no 11.340, de 7 de agosto de 2006*. http://www.planalto.gov.br/ccivil_03/_ato2004-2006/2006/lei/l11340.htm.

Bruschini, Maria Cristina A., Danielle Ardaillon and Sandra G. Unbehaum Ridenti. 1998. *Tesauro para estudos de gênero e sobre mulheres*. São Paulo, SP, Brasil: Fundação Carlos Chagas.

Bueno, Samira, Juliana Martins, Juliana Brandão, Isabela Sobral and Amanda Lagreca. 2023. "Visível e Invisível: A Vitimização de Mulheres no Brasil". 4th ed. https://forumseguranca.org.br/publicacoes_posts/visivel-e-invisivel-a-vitimizacao-de-mulheres-no-brasil-4a-edicao/.

Campos, Maria Luiza De Almeida. 2010. "O papel das definições na pesquisa em ontologia". *Perspectivas em Ciência da Informação* 15, no. 1: 220–38. https://doi.org/10.1590/S1413-99362010000100013

Campos, Maria Luiza de Almeida. 2017. "Elaboração de modelos de domínio em ontologias: a abordagem onomasiológica e a função da definição". *Ciência da Informação* 46, no. 1: 89–101. https://doi.org/10.18225/ci.inf.v46i1.4016

Campos, Maria Luiza de Almeida, Linair Maria Campos and Jackson da Silva Medeiros. 2011. "A representação de domínios de conhecimento e uma teoria de representação: a ontologia de fundamentação". *Informação & Informação* 16, no. 2: 140–64. https://doi.org/10.5433/1981-8920.2011v16nesp.p140

Conselho Nacional de Justiça. 2020. Resolução Conjunta n.º 5, de 3 de março de 2020. Institui o Formulário Nacional de Avaliação de Risco no âmbito do Poder Judiciário e do Ministério Público e dá outras providências. *DJe/CNJ* ,n.º 49 de 4 de março de 2020: 2–8. https://atos.cnj.jus.br/atos/detalhar/3218.

Dahlberg, Ingetraut. 1978. "Teoria do conceito". *Ciência da Informação* 7, no. 2: 101–7. https://doi.org/10.18225/ci.inf.v7i2.115

Dahlberg, Ingertrud. 1983; "Terminological definitions: characteristics and demands". 1983. In *Problèmes de la définition et de la synonymie en terminologie: Actes du Colloque International de Terminologie* 23-27 May 1982 Québec, Canada. Québec: Université Laval, 13-34.

EMBL-EBI. 2022. "GSSO – the Gender, Sex, and Sexual Orientation Ontology". *Repository. OLS Ontology Search* (website). https://www.ebi.ac.uk/ols4/ontologies/gsso

Gomes, Hagar Espanha, Maria Luiza de Almeida Campos and Ludmila dos Santos Guimarães. 2010. "Organização Da Informação e Terminologia: a abordagem onomasiológica". *DataGramaZero* 11, no. 5. https://cip.brapci.inf.br/download/45551

Gomes, Hagar Espanha and Maria Luiza de Almeida Campos. 2019. *A organização do conhecimento na Web: contribuições de Shiyali Ramamrita Ranganathan e de Ingetraut Dahlberg*. Série Cadernos Acadêmicos 1. Niterói: IACS/UFF.

Isotani, Seiji, and Ig Ibert Bittencourt. 2015. *Dados abertos conectados: em busca da web do conhecimento*. São Paulo: Novatec.

Legislação sobre a violência contra as mulheres no Brasil. (2023). Compromissoeaatidtude.org. http://www.compromissoeatitude.org.br/legislacao-sobre-violencia-contra-as-mulheres-no-brasil/

Mediavilla Herreros, Marisa and Rocarda Folla Fernández. *Tesauro de Mujeres*. 2014. Madrid: Biblioteca de Mujeres.

Meneghel, Stela Nazareth and Ana Paula Portella. 2017. "Feminicídios: conceitos, tipos e cenários". *Ciência & Saúde Coletiva* 22, no. 9: 3077–86. https://doi.org/10.1590/1413-81232017229.11412017

Morais, B. R. S. 2022. "Proposta de Tesauro para dar suporte a organização da informação de apoio à prevenção e combate à violência contra a mulher". PhD diss., Universidade de Brasília

Moreira, Marcelo da Silva, and Enrique Muriel-Torrado. 2019. "A relação da inteligência policial com a Ciência da Informação", In *Aproximação entre a Ciência da Informação com a Ciência Policial*., edited by Adilson Luiz Pinto, 7–19. Florianópolis: Senac SC.

Nações Unidas. 2021. "OMS: uma em cada 3 mulheres em todo o mundo sofre violência". *Nações Unidas Brasil*, March 10 2021. https://brasil.un.org/pt-br/115652-oms-uma-em-cada-3-mulheres-em-todo-o-mundo-sofre-viol%C3%AAncia.

Nações Unidas. 2022. "Objetivos de Desenvolvimento Sustentável". *Nações Unidas Brasil* (website). https://brasil.un.org/pt-br/sdgs.

ONU Mulheres. 2016. "Diretrizes Nacionais Feminicídio: Investigar, processar e Julgar com perspectiva de gênero as mortes violentas de mulheres". *Onu Mulheres*. https://www.onumulheres.org.br/wp-content/uploads/2016/04/diretrizes_feminicidio.pdf

Santos, Luana Carla de Moura dos and Marisa Bräscher. 2017. "Uso de ontologia na recuperação da informação em acervos digitais de jornais". *Informação & Informação* 22, no. 3: 346-376. https://doi.org/10.5433/1981-8920.2017v22n3p346.

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

Appendix–
Glossary OntoVDFcM

Femicide: Qualified homicide against a woman due to her female sex status; (Item VI – added by Law No. 13,104, of 3/9/2015) § 2 – A, it is considered that there are reasons for female sex when the crime involves: I – domestic and family violence; II – disregard or discrimination against the condition of women. Source: Decree Law No. 2,848, of December 7, 1940 – Brazilian Penal Code

Homicide of a woman: Killing someone – qualified homicide: if the homicide is committed: I – against payment or promise of reward, or for another base reason; II – for futile reasons; III – with the use of poison, fire, explosives, asphyxiation, torture or other insidious or cruel means, or which may result in common danger; IV – betrayal, by ambush, or through dissimulation or other resource that makes the defense of the offended party difficult or impossible; V – to ensure the execution, concealment, impunity or advantage of another crime; Source: Decree Law No. 2,848, of December 7, 1940 – Brazilian Penal Code, Section VI – added by Law No. 13,104, of 3/9/2015

Domestic and family violence concerns women: any action or omission based on gender that causes death, injury, physical, sexual or psychological suffering and moral or property damage: (See Supplementary Law No. 150, of 2015); I – within the scope of the domestic unit, understood as the space for permanent coexistence of people, with or without family ties, including those sporadically aggregated; II – within the scope of the family, understood as the community formed by individuals who are or consider themselves related, united by natural ties, affinity or express will; III – in any intimate relationship of affection, in which the aggressor lives or has lived with the offended party, regardless of cohabitation. Single paragraph. The personal relationships set out in this article are independent of sexual orientation. Art. 6 Domestic and family violence against women constitutes one of the forms of violation of human rights. Types: Sexual violence, psychological violence, moral violence, property violence, physical violence. Source: Law No. 11,340 of August 7, 2006 – Maria da Penha Law

Violence against women: any action or conduct, based on gender, that causes death, harm or physical, sexual or psychological suffering to women, both in the public and private spheres. It will be understood that violence against women includes physical, sexual and psychological violence: §1. That occurred within the family or domestic unit or in any other interpersonal relationship, in

which the aggressor lives or has lived in the same household as the woman and which includes, among others, rape, rape, mistreatment and sexual abuse: §2. That occurred in the community and is perpetrated by any person and that includes, among others, rape, sexual abuse, torture, mistreatment of people, trafficking in women, forced prostitution, kidnapping and sexual harassment in the workplace, as well as in institutions educational institutions, health establishments or any other place, and §3. That is perpetrated or tolerated by the State or its agents, wherever it occurs. Source: Decree nº 1973 of August 1, 1996.

Domestic violence: If the injury is committed against an ascendant, descendant, brother, spouse or partner, or with whom he lives or has lived, or, even, if the agent of domestic relations, cohabitation or hospitality prevails Source: Decree Law no. 2,848, of December 7, 1940 – Brazilian Penal Code

Physical violence: physical violence, understood as any conduct that offends your integrity or bodily health; Types of physical violence: Drowning; Attacks with the hands; Aggressions with objects; Threats to children; Death threats; Threats of suicide Threats of physical violence; Throwing objects; Lack of communication as a form of punishment: does not listen or does not speak; Pinching; Kicks; Decide for her; Destruction of objects with special sentimental value Pushing; Beating; Strangulation; Injuries caused by burns; Injuries caused by firearms; Injuries caused by knives or sharp objects; Shake and squeeze your arms; Prevents or hinders access to work, study or any other activity; Injuries with sharp objects; Injuries with piercing objects; Abuse of domestic animals; Bites; Sudden and unjustified changes in mood, when faced with the same situation or behavior: either praises or humiliates; Does not allow the victim to make decisions about his or her life or family, nor participate in decisions; Pauladas; Deprivation of basic needs (food, sleep, etc.); Hair pulling; Feeling of guilt and confusion, caused by constant corrections and demonstrations made by those who mistreat, with the intention of presenting themselves as victims; Punches; Suffocation; Hanging; Tapas; Torture. Source: Law No. 11,340 of August 7, 2006 – Maria da Pe Law.

Psychological violence: psychological violence, understood as any conduct that causes emotional harm and reduced self-esteem or that harms and disrupts full development or that aims to degrade or control your actions, behaviors, beliefs and decisions, through threat, embarrassment, humiliation , manipulation, isolation, constant surveillance, persistent persecution, insult, blackmail, violation of privacy, ridicule, exploitation and limitation of the right to come and go or any other means that causes harm to psychological health and self-determination; (Wording given by Law No. 13,772, of 2018). Types of psychological violence: Alienation; Abandonment; Threats; Intimidating physical approach; knocking on doors; Bribery; Embarrassment; distorting and omitting facts to leave the woman in doubt about her memory and sanity (gaslighting); examine her drawers and belongings. Exploration; Screams; Humiliation; Insistence on considering the victim crazy, stupid or useless; Insults; Isolation; Limitation of the right to come and go or any other means that causes harm to psychological health and self-determination; Insults; Express continuous jealousy and suspicion; Manipulation; Persecution; Prohibit studying; Prohibit travel; Prohibit talking to friends; Prohibit speaking to relatives; Prohibit leaving the house; Ridicule; Take away freedom of belief; Constant vigilance. Source: Law No. 11,340 of August 7, 2006 – Maria da Penha Law.

Property violence: property violence, understood as any conduct that constitutes retention, subtraction, partial or total destruction of your objects, work instruments, personal documents, assets, values and rights or economic resources, including those intended to satisfy your needs; The following types of property violence are: Purposefully causing damage to objects belonging to the woman or objects that she likes; Control the money; Control over work and salary (may retain money); Failure to pay alimony; Destruction of personal documents; Fraud; Theft, extortion or damage; Not providing access to the money necessary to meet the family's needs; Deprive of goods, values or economic resources; Theft or destruction of goods, objects, personal documents, assets, valuables, resources, etc., belonging to the woman, to cause harm and control her. Source: Law No. 11,340 of August 7, 2006 – Maria da Penha Law.

Moral violence: understood as any conduct that constitutes slander, defamation or insult. The following are types of moral violence: Accusing the woman of cheating; reports the victim to the police; devaluing the victim due to the way they dress; distorting and omitting facts to leave the woman in doubt about her memory and sanity; make moral judgments about conduct; expose intimate life; make false criticisms; Plays the victim in public, saying that she mistreats him; prevents or hinders relationships outside the couple; prevents you from accompanying you in activities or imposes your presence by force; Social isolation; demeaning the woman through insults that affect her nature; trying to tarnish the woman's reputation. Source: Law No. 11,340 of August 7, 2006 – Maria da Penha Law.

Sexual violence: understood as any conduct that forces you to witness, maintain or participate in unwanted sexual relations, through intimidation, threat, coercion or use of force; that induces her to commercialize or use, in any

way, her sexuality, that prevents her from using any contraceptive method or that forces her into marriage, pregnancy, abortion or prostitution, through coercion, blackmail, bribery or manipulation; or that limits or nullifies the exercise of your sexual and reproductive rights; The following types of sexual violence are: Coercion to maintain sexual relations, using physical force or emotional blackmail; Rape; Force marriage, pregnancy or prostitution through coercion, blackmail, bribery or manipulation; Humiliation related to her sexual conduct; Prevent the use of contraceptives or force women to have an abortion; Limit or annul the exercise of women's sexual and reproductive rights; Forcing women to sell or use, in any way, their sexuality; Forcing a woman to perform sexual acts that cause discomfort or repulsion; Degrading treatment of females; Violence and aggression during pregnancy. Source: Law No. 11,340 of August 7, 2006 – Maria da Penha Law.

Drag Storytimes and Bibliographic Invisibility: A Comparative Analysis of Picture Book Subject Metadata

Sarah Barriage*, Beth Strickland Bloch**, and Vanessa Kitzie***

* **320 Little Library Building, School of Information Science,
University of Kentucky, Lexington, KY 40502 USA
*** 1501 Greene Street, Room 211, School of Information Science,
University of South Carolina, Columbia, SC 29208 USA
*sarah.barriage@uky.edu, **beth.S.Bloch@uky.edu, ***kitzie@mailbox.sc.edu

Sarah Barriage is an Associate Professor in the School of Information Science at the University of Kentucky. She completed her PhD in communication, information, and library studies in the School of Communication & Information at Rutgers, The State University of New Jersey. Dr. Barriage's research interests include the information practices of children and youth, social justice in information institutions, and the development and use of child-centered research methods. Her work has been published in scholarly journals such as *Library & Information Science Research*, *Library Quarterly*, *International Journal of Qualitative Methods*, and *Journal of Childhood Studies*.

Beth Strickland Bloch is an Assistant Professor in the School of Information Science at the University of Kentucky. She completed her PhD in library and information science in the School of Information Sciences at the University of Illinois, Urbana-Champaign. Dr. Bloch's research interests include the values and ethics associated with the construction of professional identities, the design of technological systems, and the organization of information. Her work has been published in scholarly venues including *International Journal of Qualitative Methods* and the Association for Information Science and Technology conference proceedings.

Vanessa Kitzie is an Associate Professor in the School of Information Science at the University of South Carolina. She completed her PhD in communication, information, and library studies in the School of Communication & Information at Rutgers, The State University of New Jersey. Dr. Kitzie's research examines the information practices of LGBTQIA+ populations and how information centers and systems can better serve them. Her research has been published in journals such as *Library & Information Science Research*, *Journal of the Association for Information Science and Technology*, and *Journal of Documentation*.

Barriage, Sarah, Beth Strickland Bloch, and Vanessa Kitzie. 2024. "Drag Storytimes and Bibliographic Invisibility: A Comparative Analysis of Picture Book Subject Metadata". *Knowledge Organization* 51, no. 8: 686-699. 58 references. DOI:10.5771/0943-7444-2024-8-686.

Abstract: Historically, library materials about diverse identities have often been subject to what Gough and Greenblatt (1992) term "systemic bibliographic invisibility," the use of "outmoded, prejudicial, inadequate, or inappropriate terminology" (61) within bibliographic records to describe an item's contents. Using such terminology within subject metadata can make materials challenging to find within a library's catalog, restricting users' access to the materials and the ideas they contain. Prior work has demonstrated that folksonomies like LibraryThing may better represent the multiplicity and fluidity of marginalized identities. In this study, we analyze the subject metadata associated with a corpus of picture books read during drag storytimes, comparing the inclusion of different types of subject metadata found in bibliographic records from the Library of Congress catalog and LibraryThing. Specifically, we analyze the use of terms that explicitly describe various facets of human difference and those that refer to diverse elements within the books in more generalized or implicit terms within the bibliographic records of picture books that include depictions of LGBTQIA+ characters and/or themes, BIPOC characters, and characters with disabilities, developmental differences, and chronic illnesses. LibraryThing records contained a higher prevalence of subject metadata types across nearly all book categories, indicating that users assign more of a variety of types of subject metadata than do professional catalogers. Implications for the discoverability and accessibility of children's materials depicting marginalized identities are discussed.

Received: 29 June 2024; Accepted 31 July 2024.

Keywords: subject metadata; marginalized identities; bibliographic records.

1.0 Introduction

Diversity in children's literature is essential in addressing the biases and prejudices that can develop in early childhood (Nel 2017; Brown 2021). As the oft-quoted Bishop (1990) aptly stated, books can function as mirrors, windows, and sliding glass doors, allowing children to see themselves reflected in the stories they read, gain insight into the experiences of others who are different from themselves, and develop empathy and a change in perspective.

Incorporating diverse books within storytime programs is one way that children's librarians can work to address bias and oppression by offering programs that reflect the diversity of today's world, part of the critical competencies of the profession (Association for Library Service to Children 2020). As Bratt (2022, 26) argues: "We librarians can be the ones, through our storytimes, to show the vast array of humanity, normalizing all people and talking positively about difference. Librarians have the power to make different people or experiences either visible or invisible in the storytime space".

There is a growing body of evidence indicating that drag storytimes include representations of diverse identities in the books chosen for these programs, including characters of various gender identities, sexualities, races/ethnicities, and disability statuses (e.g., Barriage et al. 2024; Naidoo 2018). In particular, books read at drag storytimes often explicitly focus on LGBTQIA+ (lesbian, gay, bisexual, trans, queer, intersex, asexual, and more) identities and themes to both visibilize and normalize them, which offers children the opportunity to see identities, families, and experiences that may look like their own or like those of others around them (Barriage et al. 2024; Naidoo 2018).

Historically, library materials about diverse identities have often been subject to what Gough and Greenblatt (1992, 61) term "systemic bibliographic invisibility," the use of "outmoded, prejudicial, inadequate, or inappropriate terminology" within bibliographic records to describe an item's contents. Using such terminology within subject metadata can make these materials challenging to find within a library's catalog, restricting users' access to the materials and the ideas they contain. Scholars concerned with the bibliographic invisibility of materials related to race/ethnicity (e.g., Strottman 2007; Snow and Dunbar 2023), disability (e.g., Johnson and Forsythe 2019; Watson and Schaefer 2023), and LGBTQIA+ identities (e.g., McClary and Howard 2007; Adler 2009), have primarily focused on subject access related to adult and/or young adult library materials. In this study, we focus on materials for children – specifically, we analyze the subject metadata associated with picture books featuring diverse identities and/or themes read during drag storytimes, focusing on gender identity, sexual orientation, gender roles/stereotypes/norms, race/ethnicity, and disabilities, development differences, and chronic illnesses (DDDCI).

2.0 Literature review

2.1 Drag storytimes

Drag storytimes are children's events that feature drag performers (including drag queens, drag kings, and non-binary and gender non-conforming performers, among other forms of drag artistry) reading children's books and engaging in other storytime activities. These events were initiated in 2015 in Canada as Drag Queen StoryTime, hosted by Reelout Arts Project Inc. (https://www.reelout.com/about/dragqueenstorytime/), and in the United States as Drag Queen Story Hour (now named Drag Story Hour [DSH]), founded by Michelle Tea and queer literary-arts organization RADAR Productions (https://www.dragstoryhour.org/about). DSH has developed into a network of self-managed and financed chapters in the United States and internationally (Montague and Latham 2019). However, drag storytime programming does not always follow the DSH model. Many libraries and other institutions (e.g., community spaces, bookstores) host such programs independently or in collaboration with LGBTQIA+ organizations within their communities (Barriage et al. 2021; Naidoo 2018).

Public libraries and other institutions hosting drag storytimes often promote them as general inclusivity programming (Naidoo 2018). Such promotion focuses on the benefits of storytimes for children and families, most of whom are not LGBTQIA+. Benefits evidenced in the research literature include increasing family interactions by facilitating open and honest communication and exposing attendees to gender-expansive concepts (Montague and Latham 2019; Radis et al. 2022), which can lead to increased knowledge and acceptance of diverse gender identities and expressions (Radis et al. 2022). Some promotion of storytimes focuses on rainbow families (families with LGBTQIA+ parents and/or children) and gender-diverse children specifically (Naidoo 2018). Benefits to these audiences include rainbow families feeling seen and included and gender-diverse children feeling supported in their identities, which can lead to better mental health outcomes (Westwater et al. 2019). Drag storytimes may also be framed as programs that focus on messages related to human difference, acceptance, and inclusion more generally (Kitzie et al. 2022; Staino 2017; Radis et al. 2020).

Many drag storytime performers have experience and training in early childhood education (Barriage et al. 2021; Kitzie et al. 2022; Montague and Latham 2019). They wield this experience to significant pedagogical effect, as evidenced in publications authored by performers that outline the dramaturgical and pedagogical strategies and benefits

inherent in drag as an art form (Dorsey 2020; Keenan and Lil Miss Hot Mess 2020). A key benefit of storytimes is imparting critical literacy skills that allow children to interrogate gender and other social structures that promote "normalcy", which enhances their engagement, enjoyment, and literacy development during storytimes (Sipe 2008).

In addition to the art of drag performance and storytime activities, the books performers read are a critical event component. Many public library staff members who have hosted drag storytimes report using children's books with LGBTQIA+ and/or gender-non-conforming characters (Barriage et al. 2024; Naidoo 2018). For some children, these programs may be the only instances of library programming in which they are exposed to books that feature LGBTQIA+ characters and themes. For example, a recent survey found that less than half of public library respondents reported offering LGBTQIA+ programming (e.g., Pride storytimes) other than drag storytimes (Naidoo 2018), while an analysis of the picture books read during public library storytimes noted that only one book out of the corpus of 160 featured characters who could be characterized as LGBTQIA+ (Cahill et al. 2021). Public library staff also report incorporating books in drag storytimes that feature various diverse identities in addition to LGBTQIA+ characters, such as characters of various races/ethnicities and characters with DDDCI (Barriage et al. 2024), aligning with the aim of promoting diversity and inclusion more generally.

However, not all books read at drag storytimes include diverse identities and themes (Barriage et al. 2024). Research has consistently shown that there is a lack of diversity in the picture books published each year (Larrick 1965; Adukia et al. 2023), which is likely a contributing factor. Another factor that may contribute to the lack of diverse books in drag storytimes is that such books only sometimes align with the best read-aloud practices. For example, one library director stated: "Some of the books about diverse gender expression can be long and we have an audience of very young children at these programs" (Naidoo 2018, 19). To help mitigate these factors, library staff frequently work with drag performers to select books for use in drag storytime, providing recommendations of specific books that work well as read-alouds and fit within the storytime's theme and/or guidelines for making book selections (Barriage et al. 2021; Condren 2018; Naidoo 2018); however, the problem persists.

2.2 Subject access and metadata

One strategy that library staff, drag performers, and patrons alike may use to identify picture books with diverse identities and/or themes is through subject searches in library catalogs. Subject headings, such as Library of Congress Subject Headings (LCSH) and Children's Subject Headings (CSH), provide standardized terms that describe the subject matter of a particular item, and librarians will assign these to items while conducting a subject analysis during the cataloging process (Hoffman 2019). Subject headings provide an access point for users when looking for information in various systems (Joudrey and Taylor 2018). Yet, determining the "aboutness" of an item is a subjective conceptual activity (Hauser and Tennis 2019), one that is often taken for granted and rarely critically examined (Holley and Joudrey 2021). Scholars who question how subject headings are developed and applied argue that their seeming neutrality often hides sociopolitical biases (Olson 2001; Drabinski 2013). People depend on a variety of values (i.e., functional, social, moral) when designing systems (Nissenbaum 2001; Friedman et al. 2013), and these values translate into the tools and structures they develop (Winner 1980). Depending on how these values are expressed, either explicitly or implicitly, will contribute to the fallible representation, or often lack of representation, of marginalized identities in library catalogs and controlled vocabularies. The failure to adequately include diverse identities in surrogate records results in information becoming "either unfindable or unusable by members of marginalized groups" (Dobreski et al. 2022, 490-491).

Since users rely on information infrastructures to guide their information-seeking process, this dependence results in using controlled classification systems which hinder adequate knowledge production (Bowker and Star 2000; Cifor and Rawson 2023). Yet, not all contemporary information systems assign subject access points using a controlled vocabulary.

For instance, the online social cataloging platform LibraryThing allows users to apply their own tags to items. Tags are composed of single words or phrases that users apply to items to describe their content (Rolla 2009). This kind of uncontrolled vocabulary provides a way to move beyond traditional subject classifications based on literary warrant to ones based on 'user warrant' and the language of the end-user (Moulaison and Bossaller 2017). Comparative studies of applied LCSH and LibraryThing tags have found that there tends to be a disconnect between the two schemes when used to classify LGBTQIA+ materials (Adler 2009; Rolla 2009) and materials about ethnic minorities (Bates and Rowley 2011). Similar discrepancies have been observed when comparing the language used to describe disabilities in LCSH and user-generated tags on Archive of Our Own (Johnson and Forsythe 2019). Although the reasons for these misalignments need further evaluation, one reason might relate to what Wagner (2022, xii) argues is catalogers' hesitation in describing humans in surrogate records because of "given societal complexities around identities". This observation, in part, reflects how the political and ethical landscapes associated with cataloging practices ultimately lead to the increased invisibility of marginalized identities.

Although subject headings and tags have received much attention in the scholarly literature, little work has focused explicitly on subject metadata assigned to children's literature. One example of such work is that by Mendell and Sarles (2010), who found that LCSH at that time did not include any subject headings to describe donor conceived people (e.g., children conceived via egg and/or sperm donation). As Mendell and Sarles (2010) note, this made finding books on this specific topic quite challenging, demonstrating the "systemic bibliographic invisibility" of marginalized materials described by Gough and Greenblatt (1992, 61).

More recently, Williams (2017) analyzed the subject headings assigned to books for children and young adults with diversity-related content (specifically, books focused on topics related to race/ethnicity, LGBTQIA+ identities, and disability/illness). Nearly 86% of the 120 books analyzed in this study included at least one subject heading explicitly related to its diverse content. However, a higher proportion of the LGBTQIA+ books did not have at least one explicitly diverse subject heading in their catalog records than the records of books focused on race/ethnicity and disability/illness. Williams (2017, 17) noted that some of the catalog records for the LGBTQIA+ books instead contained what they termed "fluff" subject headings, "ambiguous or seemingly meaningless" subject headings that do not explicitly communicate the diverse content. Williams (2017) suggests that applying such 'fluffy' subject headings instead of those that explicitly reflect a book's diverse content may reflect a move towards normalizing LGTBQIA+ content; however, it may also make these books more challenging to find in a library's catalog.

3.0 Current study

This comparative study examines the presence of explicitly diverse and 'fluffy'[1] subject metadata within the metadata records of a corpus of picture books read during drag storytimes featuring diverse identities and/or themes in the Library of Congress Catalog (LCC) and on the LibraryThing platform. Within this corpus, our focus is on books with characters/themes related to gender identity, sexual orientation, and gender roles/norms/stereotypes (combined within the LGBTQIA+ category), characters who are Black, Indigenous, and people of color (BIPOC), and characters with DDDCI. Specifically, our analysis aims to answer the following:

RQ1a: What frequency/percentage of picture books with LGBTQIA+, BIPOC, and DDDCI characters and/or themes read during drag storytimes have explicitly diverse, fluffy, and other subject headings assigned to their bibliographic records in LCC?

RQ1b: How do these percentages differ based on book type?

RQ2a: What frequency/percentage of picture books with LGBTQIA+, BIPOC, and DDDCI characters and/or themes read during drag storytimes have explicitly diverse tags, fluffy tags, and other tags assigned to their metadata records in LibraryThing?

RQ2b: How do these percentages differ based on book type?

RQ3a: How prevalent are explicitly diverse, fluffy, and other subject metadata for picture books with LGBTQIA+, BIPOC, and DDDCI characters and/or themes read during drag storytimes in LCC versus LibraryThing?

RQ3b: How do these proportions differ based on book type?

4.0 Methods

The analysis reported here builds on a larger, multi-phase project examining drag storytimes in public libraries. Previous phases of the study included a survey of library staff and interviews with library staff and drag performers related to their perceptions of and experiences with drag storytimes (Barriage et al. 2021; Kitzie et al. 2022; Oltmann et al. 2023), as well as a content analysis of diversity in picture books read during drag storytimes (Barriage et al. 2024).

4.1 Initial sample

An initial list of picture books read during drag storytimes in public libraries was generated by: 1) reviewing news articles and professional/scholarly literature on drag storytimes to identify specific titles of picture books read during drag storytime events; 2) reviewing the transcripts of interviews with drag performers and library staff working at libraries that have hosted drag storytimes for mention of specific titles of picture books read during drag storytime events; and 3) a brief survey of library staff who have hosted drag storytimes in the past.

This process resulted in a list of 103 picture books after removing duplicates and book titles that were either generic or did not match any record in WorldCat (a bibliographic database combining data about items in library collections worldwide; https://worldcat.org). Complete details related to the generation of this initial list are reported elsewhere (Barriage et al. 2024).

4.2 Data collection

We obtained subject headings assigned to the picture books from LCC (https://catalog.loc.gov/). As these books are children's materials, their bibliographic records contained

subject headings drawn from LCSH and/or CSH. We recorded subject headings from both/either of these two controlled vocabularies included in each book's bibliographic record into a spreadsheet. See Figure 1 for an example of a bibliographic record in LCC for one of the books in our corpus.

We accessed LibraryThing (https://www.librarything.com/) tags manually since the API has been discontinued. Once each book's metadata record was located, we used the "show all tags" and "numbers" filters, indicating the frequency with which each tag was applied. We recorded all tags for each book in a spreadsheet and then sorted tags by

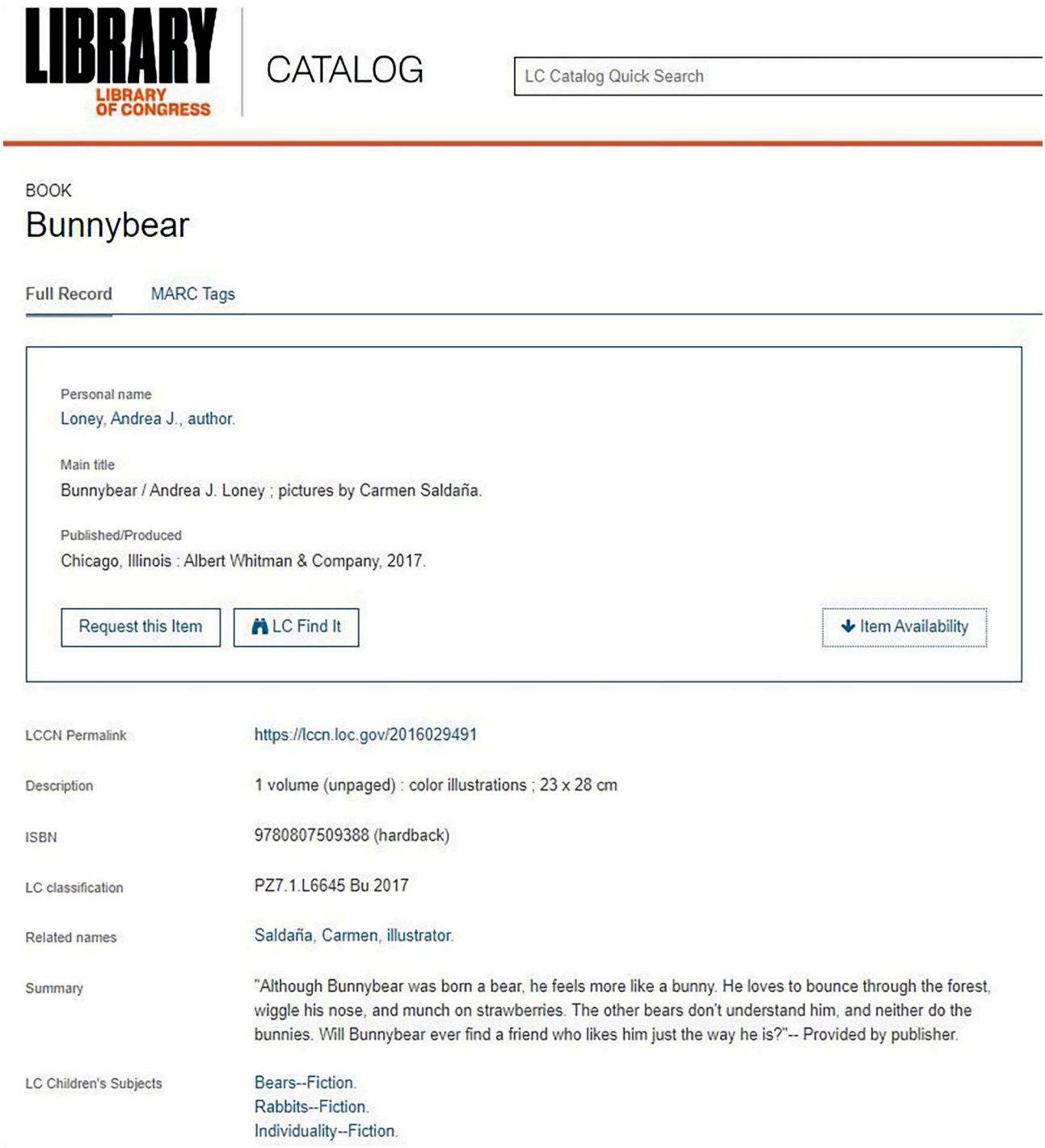


Figure 1. Partial LCC bibliographic record for *Bunnybear* by Andrea Loney, including subject metadata assigned from CSH.

frequency, keeping the top ten tags for each title as these are both the most used and the most visible tags when using the LibraryThing platform. In some cases, there were more than ten tags for a book when the tenth place had a tie (i.e., all tags that tied for the tenth place were included in our analysis). In other cases, we included fewer than ten tags as we only selected tags that two or more people had applied. While LibraryThing does include information about the frequencies of applied tags, we did not include that information within this analysis. See Figure 2 for an example of a LibraryThing record for one of the books in our corpus.

Two books did not have records in either LCC or LibraryThing and were thus excluded from further analysis, leaving a total of 101 picture books that had records in LCC and/or LibraryThing.

4.3 Data analysis

Once all subject headings and tags were collected for the final list of 101 picture books, these subject metadata were coded for the presence of subject headings/tags that were explicitly about topics and/or themes related to LGBTQIA+ identities (coded as "explicit - LGBTQIA+"; for example, the subject heading "Gay parents"), race/ethnicity (coded as "explicit - race/ethnicity"; for example, the LibraryThing tag "African American"), or DDDCI (coded as "explicit - DDDCI"; for example, the subject heading "People with disabilities"), those that were more generalized and/or implicit (coded as "fluffy"; for example, the LibraryThing tag "inclusion"), and those that were unrelated to these diverse topics and/or themes more generally (coded as "other"; for example, the subject heading "Board books"). Two research team members independently coded the subject headings and tags for each book. The research team then met to discuss all discrepancies.

Using the results of the prior quantitative content analysis of the picture books in our sample (Barriage et al. 2024), we then coded for the presence of the following in each book (note that these categories are not mutually exclusive):

- Depiction of LGBTQIA+ lead and/or non-lead characters (e.g., *I Am Jazz* by Jessica Herthel and Jazz Jennings 2014);
- Themes related to gender roles/stereotypes/norms (e.g., *Morris Micklewhite and the Tangerine Dress* by Christine Baldacchinoa 2014);
- Themes with queer subtext (e.g., *BunnyBear* by Andrea Loney 2021); [2]
- Depiction of BIPOC lead and/or non-lead characters (e.g., *We March* by Shane W. Evans 2016); and/or
- Depiction of lead and/or non lead characters with a DDDCI (e.g., *We're All Wonders* by R. J. Palacio 2017).

After we resolved coding discrepancies, we downloaded the coding results into Excel and cleaned the data, which involved converting all counts of LCC and LibraryThing subject metadata from a sum (e.g., four total explicit LCC subject headings for a certain book) to binary values where 1 signified that there was at least one subject heading/tag of that type and 0 signified that there were no subject headings/tags of that type. We then converted all book categories (e.g., LGBTQIA+ lead) to binary, categorical values (e.g., LGBTQIA+ lead, no LGBTQIA+ lead) to facilitate building of pivot tables. We then generated pivot tables that obtained the frequencies of books in various categories (e.g., LGBTQIA+ lead) with different types of subject metadata (i.e., explicit, fluffy, or other) divided by whether the metadata were from LCC or LibraryThing. We also generated tables tabulating the total number of books within a category, dividing the frequency of books with different subject metadata by total number of books in that category. This division gave the percentage of books in each category with a specific type of subject metadata. We then calculated the absolute value of the difference in percentages between LibraryThing and LCC.

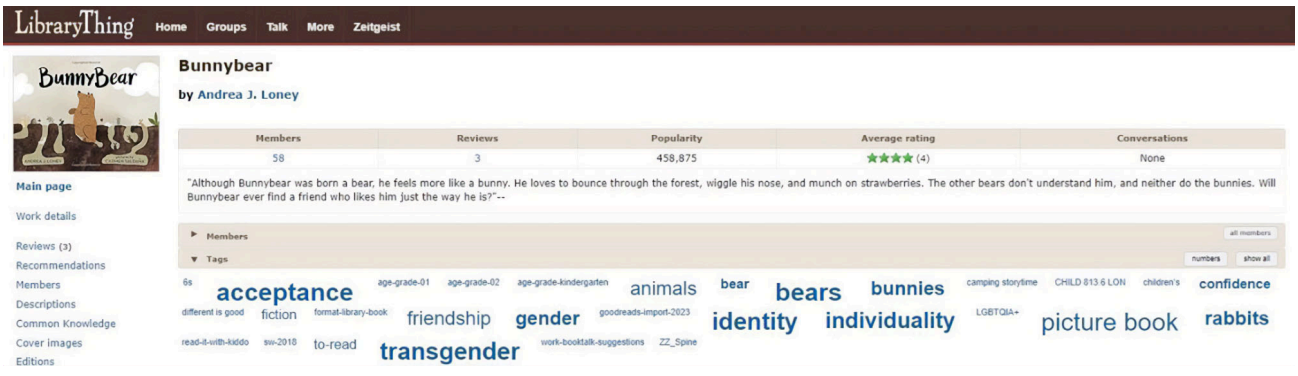


Figure 2. Partial LibraryThing metadata record for *Bunnybear* by Andrea Loney, including user-generated subject metadata.

5.0 Results

Ninety-two books had records in the LCC. Overall, thirty-nine books had explicit subject headings, thirty-two had fluffy subject headings, and ninety had other subject headings. Ninety-nine books had records in LibraryThing. Fifty-three books had explicit tags, fifty-eight had fluffy tags, and ninety-nine had other tags. Figure 3 shows the percentages of all books included in this analysis with each type of subject metadata in LCC and LibraryThing.

5.1 Picture books with LGBTQIA+ characters and/or themes

Forty-seven percent (n = 47) of all books reviewed (N = 101) had lead and/or non-lead LGBTQIA+ characters, themes related to gender roles/stereotypes/norms, and/or queer subtext. Eleven percent (n = 11) of books had an LGBTQIA+ lead character, eighteen percent had an LGBTQIA+ non-lead character (n = 18), seventeen percent (n = 17) included themes related to gender roles/stereotypes/norms, and eleven percent (n = 6) had themes with queer subtext. Across all subject metadata types (explicit, fluffy, and other), the frequency of LibraryThing tags exceeded the frequency of LCC subject headings (see Figure 4).

5.1.1 Explicit subject metadata

Nineteen percent ($n_{LibThing}$ = 38, 81%; n_{LCC} = 29, 62%) more books with LGBTQIA+ characters, themes related to gender roles/stereotypes/norms, and/or queer subtext had explicit subject metadata in LibraryThing as compared to LCC. The largest difference in favor of explicit LibraryThing subject metadata was in books with queer subtext ($n_{LibThing}$ = 3, 50%; n_{LCC} = 0, 0%), followed by books related to gender roles/stereotypes/norms ($n_{LibThing}$ = 16, 94%; n_{LCC} = 13, 76%). There were also more books with explicit subject metadata in LibraryThing among books with LGBTQIA+ non-leads ($n_{LibThing}$ = 14, 78%; n_{LCC} = 12, 67%) and LGBTQIA+ leads ($n_{LibThing}$ = 10, 91%; n_{LCC} =9, 82%).

5.1.2 Fluffy subject metadata

The differences between LCC and LibraryThing were particularly marked when considering fluffy subject metadata. Here, 45% ($n_{LibThing}$ = 38, 81%; n_{LCC} = 17, 36%) more books with LGBTQIA+ characters, themes related to gender roles/stereotypes/norms, and/or queer subtext had subject metadata coded as fluffy in LibraryThing as compared to LCC. The largest difference was found when comparing the difference in books with LGBTQIA+ non-lead characters assigned fluffy subject metadata ($n_{LibThing}$ = 14, 78%; n_{LCC} = 4,

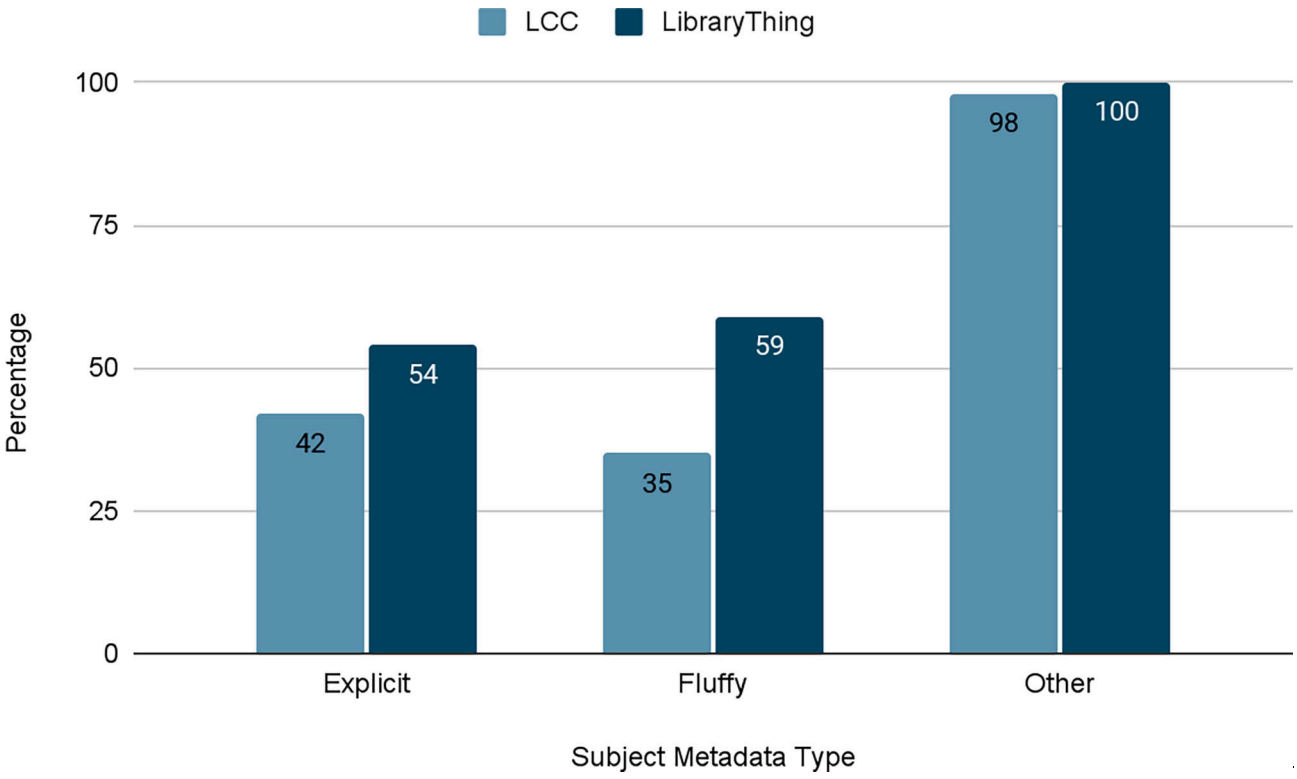


Figure 3. Percentage of books of all types with explicit, fluffy, and other subject metadata in LCC and LibraryThing.

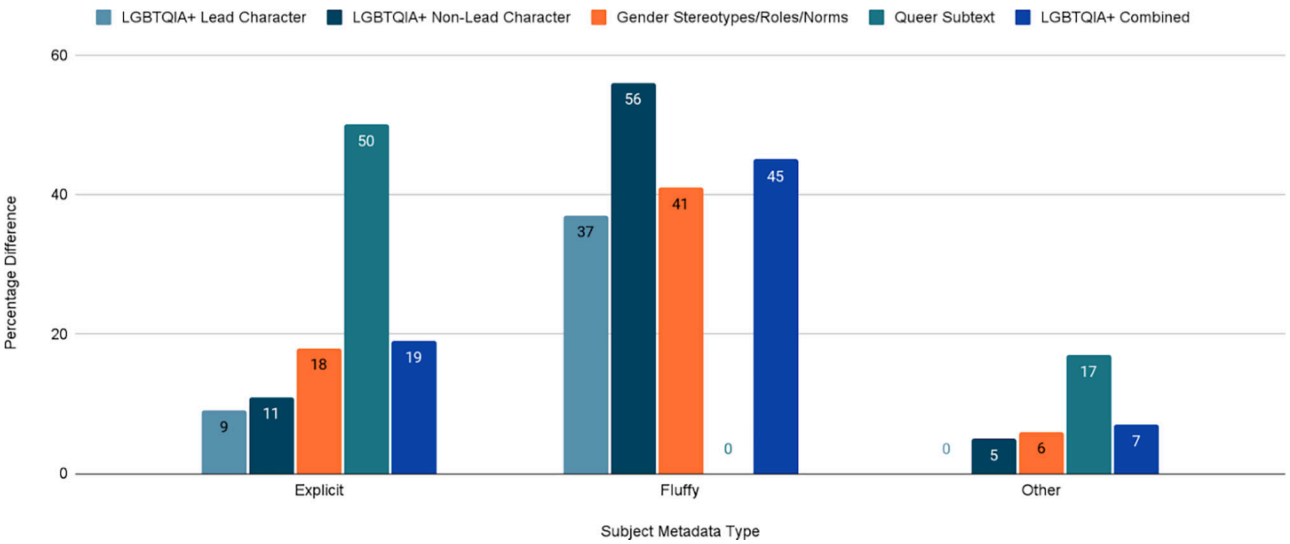


Figure 4. Difference in percentages between LibraryThing and LCC for each subject metadata type for books with LGBTQIA+ characters and/or themes. Positive values indicate LibraryThing had a greater frequency of metadata type than did LCC; negative values indicate LCC had a greater frequency of metadata type than did LibraryThing.

22%). Differences in books with LGBTQIA+ lead characters ($n_{LibThing}$ = 7, 64%; n_{LCC} = 3, 27%) and themes related to gender roles/stereotypes/norms ($n_{LibThing}$ = 13, 76%; n_{LCC} = 6, 35%) were a little less, but still pronounced. There was no difference among books with queer subtext regarding number of fluffy subject headings ($n_{LibThing}$ = 6, 100%; n_{LCC} = 6, 100%).

5.1.3. Other subject metadata

Differences in subject metadata classified as other between LibraryThing and LCC were present, but lesser in degree compared to explicit and fluffy subject metadata. Overall, there was only a 7% ($n_{LibThing}$ = 45, 96%; n_{LCC} = 42, 89%) difference in books with LGBTQIA+ characters, themes related to gender roles/stereotypes/norms, and/or queer subtext assigned other metadata in LibraryThing as compared to LCC. The largest difference was present for books with queer subtext at 17% ($n_{LibThing}$ = 6, 100%; n_{LCC} = 5, 83%). Books with themes related to gender roles/stereotypes/norms ($n_{LibThing}$ = 17, 100%, n_{LCC} = 16, 94%) had a 6% difference in frequency assigned in LibraryThing versus LC and books with LGBTQIA+ non-lead characters ($n_{LibThing}$ = 17, 94%; n_{LCC} = 16, 89%) had a 5% difference. There was no difference ($n_{LibThing}$ = 10, 91%; n_{LCC} = 10, 91%) between books with LGBTQIA+ lead characters assigned other subject metadata in LibraryThing and LCC.

5.2 Picture books with BIPOC characters

Fifty two percent (n = 53) of all books (N = 101) had a BIPOC lead character (15%, n = 15) and/or supporting character (52%, n = 53). The frequency of LibraryThing tags exceeded the frequency of LCC subject headings across all subject metadata types, with the exception of explicit subject metadata applied to books with BIPOC lead characters (see Figure 5).

5.2.1 Explicit subject metadata

The differences between LCC and LibraryThing in regards to explicit subject metadata assigned to books with BIPOC characters were small. LibraryThing records had a higher percentage of books in this category overall with explicit subject metadata than did LCC, with a difference of 2% ($n_{LibThing}$ = 8, 15%; n_{LCC} = 7, 13%). These differences are the same for books with BIPOC non-lead characters only. However, explicit LCC subject headings for books with BIPOC lead characters exceeded LibThing tags by 7% ($n_{LibThing}$ = 3, 20%; n_{LCC} = 4, 27%).

5.2.2 Fluffy subject metadata

The largest differences between the types of subject metadata applied to books with BIPOC lead and/or non-lead characters was in the fluffy category, where 36% more LibraryThing books ($n_{LibThing}$ = 39, 74%) had this type of subject metadata compared to LCC (n_{LCC} = 20, 38%). The same percentage and frequency differences in favor of LibraryThing applied to books with BIPOC non-lead characters. There was a 26% difference in favor of LibraryThing books with BIPOC lead characters containing fluffy subject metadata ($n_{LibThing}$ = 11, 73%; n_{LCC} = 7, 47%).

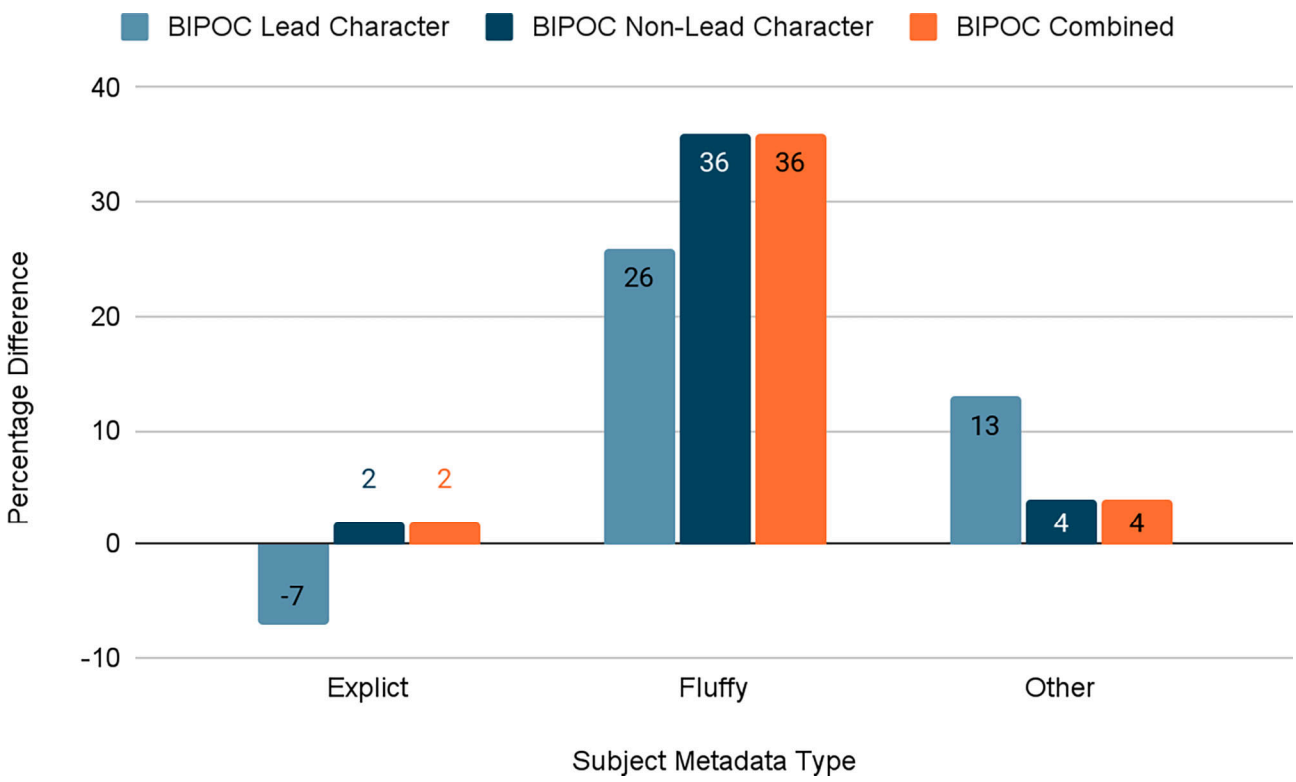


Figure 5. Difference in percentages between LibraryThing and LCC for each subject metadata type for books with BIPOC characters. Positive values indicate LibraryThing had a greater frequency of metadata type than did LCC; negative values indicate LCC had a greater frequency of metadata type than did LibraryThing.

5.2.3 Other subject metadata

Subject metadata classified as other was found in more records in LibraryThing than LCC, particularly for books featuring BIPOC leads where there was a 13% difference ($n_{LibThing}$ = 15, 100%; n_{LCC} = 13, 87%). There were differences between books with BIPOC supporting characters and all BIPOC characters combined, but this difference was very small at 4% for both ($n_{LibThing}$ = 51, 96%; n_{LCC} = 49, 92%).

5.3 Picture books with characters with a DDDCI

Fifty-seven percent (n = 58) of all books (N = 101) had a lead character (57%, n = 58) and/or a supporting character with a DDDCI (55%, n = 56). Compared to the other book types, there was the most variation between prevalence of subject metadata types among LibraryThing and LCC for books of this type (see Figure 6).

5.3.1 Explicit subject metadata

There were very slight differences in favor of LCC having more books assigned explicit subject metadata compared to LibraryThing for all books with characters with a DDDCI ($n_{LibThing}$ = 2, 3%; n_{LCC} = 3, 5%), as well as books with non-lead characters with a DDDCI ($n_{LibThing}$ = 2, 4%; n_{LCC} = 3, 5%). There were 20% more books in LCC for books with lead characters with a DDDCI assigned explicit metadata than in LibraryThing ($n_{LibThing}$ = 0, 0%; n_{LCC} = 1, 20%). However, it should be noted that these numbers are very small.

5.3.2 Fluffy subject metadata

When it came to fluffy subject metadata, all books with a character with a DDDCI in LibraryThing that were assigned fluffy tags outnumbered those with assigned fluffy subject headings in LCC at a 27% difference ($n_{LibThing}$ = 39, 67%; n_{LCC}= 23, 40%). There was a similar 29% difference favoring the number of books assigned fluffy tags in LibraryThing among books with non-lead characters with a DDDCI ($n_{LibThing}$ = 39, 70%; n_{LCC} = 23, 41%). However, there were more books with lead characters with a DDDCI assigned fluffy subject headings in LCC compared to LibraryThing tags, with a 20% difference ($n_{LibThing}$ = 1, 20%; n_{LCC} = 2, 40%) although the total number of books was very small for this category.

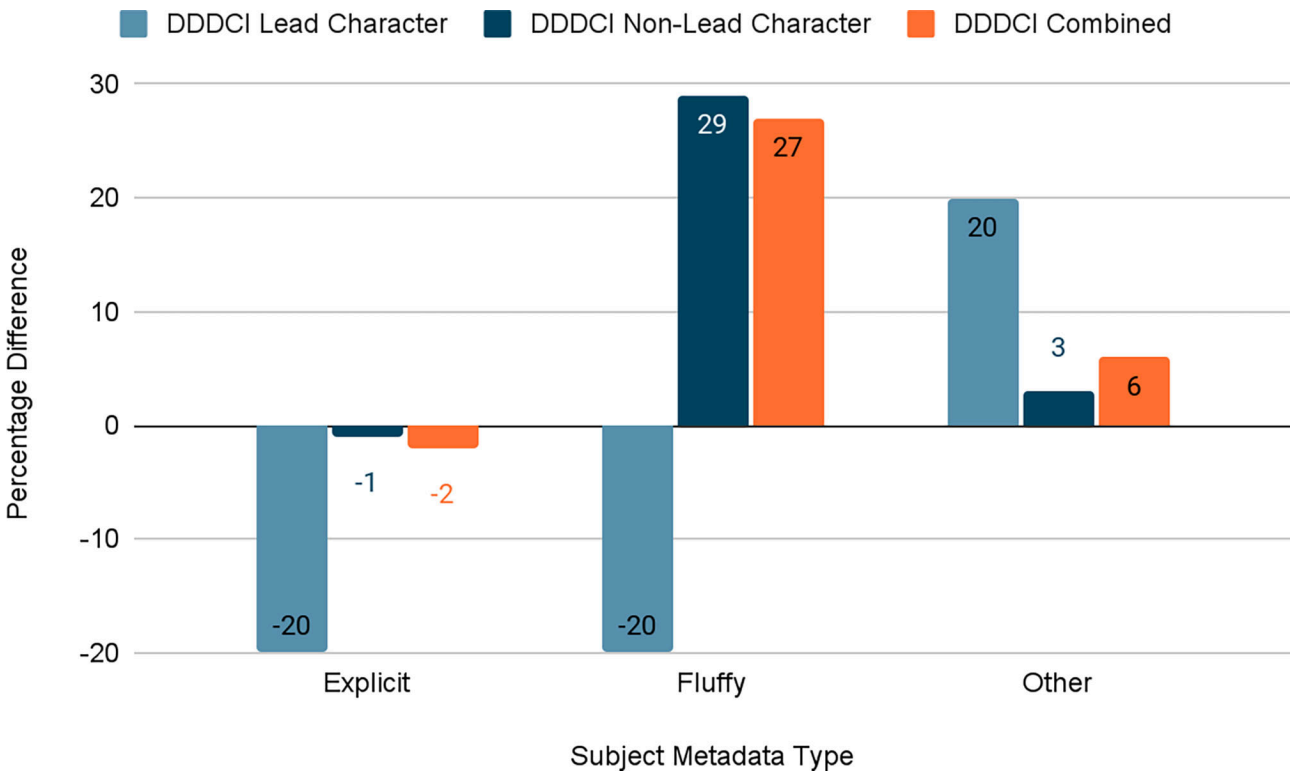


Figure 6. Difference in percentages between LibraryThing and LCC for each subject metadata type for books with characters with DDDCI. Positive values indicate LibraryThing had a greater frequency of metadata type than did LCC; negative values indicate LCC had a greater frequency of metadata type than did LibraryThing.

5.3.3 Other subject metadata

There were minimal differences slightly in favor of LibraryThing books featuring characters with a DDDCI with subject metadata coded as other at 6% (n_{LibThing} = 56, 97%; n_{LCC} = 53, 91%). This difference was similar for books with non-lead characters with a DDDCI at a 3% difference (n_{LibThing} = 54, 96%; n_{LCC} = 52, 93%). Twenty percent more books with lead characters with a DDDCI had other subject metadata in LibraryThing as compared to LCC (n_{LibThing} = 5, 100%; n_{LCC} = 4, 80%), although again the total number of books was small for this category.

6.0 Discussion

This study found that users typically assign more of a variety of subject tags across the explicit, fluffy, and other categories in LibraryThing as compared to the variety of subject heading types assigned by librarians in the LCC for books that include LGBTQIA+, BIPOC, and DDDCI characters/themes. These findings align with previous research suggesting that uncontrolled vocabularies are less restrictive than controlled vocabularies (Rolla 2009; Johnson and Forsythe 2019). Uncontrolled vocabularies prioritize 'user warrant' over the traditional 'literary warrant' to determine an item's aboutness (Moulaison and Bossaller 2017). Our findings indicate that user warrant for children's materials may be quite different from literary warrant, with different types of subject metadata being applied by users than that assigned by librarians. In some cases and contexts, 'fluffy' concepts such as inclusion or individuality found in user tags may be more salient to users of children's materials than more explicit concepts such as cultural pluralism and feminism found in subject headings. Ultimately, this underscores the potential of user-generated tags to enhance the visibility and accessibility of diverse materials in library catalogs.

The higher prevalence of explicit user tags for books with LGBTQIA+ characters may reflect Wagner's (2022) assertion that catalogers are avoiding the application of identity-based subject headings in this context. However, explicit subject metadata were more frequently applied in LCC to books with BIPOC lead characters and both lead and non-lead characters with DDDCI as compared to LibraryThing. This may suggest that catalogers are less likely to apply identity-based subject metadata related to gender identity and sexual orientation than identity-based subject metadata related to race/ethnicity and DDDCI.

Although the frequency of tag and subject heading application in general, and of explicit terms specifically, offers valuable insights, a comparative analysis of explicit and fluffy subject metadata reveals the most compelling findings. This comparison may also shed light on the differing approaches catalogers and users take in applying subject metadata. Records in LibraryThing had a higher frequency of fluffy tags as compared to fluffy subject headings in LCC for all books with LGBTQIA+ characters, books about gender stereotypes/roles/norms, books with BIPOC characters, and books with non-lead characters with DDDCI. Although the use of fluffy subject headings might be seen as a way to make items more challenging to find in a library's catalog (Williams 2017), the same might not be true for the use of fluffy user-generated tags. This discrepancy may suggest that users on platforms like LibraryThing, who often generate tags based on personal interpretations and preferences, might apply broader, more generalized tags to engage a wider audience or to reflect personal understandings of diversity that differ from the more standardized subject headings found in LCC. Additionally, the predominance of fluffy tags could be influenced by a community-driven approach to categorization, which emphasizes inclusivity and accessibility over precise terminological accuracy. This again speaks to the methodological flexibility of user-generated tags and increasingly the findability of diversity-related children's books.

Notably, all books with queer subtext were assigned fluffy subject metadata in both LCC and LibraryThing. No book with queer subtext was assigned explicit subject metadata in LCC, while half were assigned explicit tags in LibraryThing. Although fluffy subject metadata can be used to obfuscate marginalized identities when used in conjunction with explicit labels, they bring attention to LGBTQIA+ themes that may otherwise be missed. This latter claim is especially valid given that subject tagging can be used to denote minor subject themes. This finding suggests that tags may be better at identifying LGBTQIA+ themes in children's literature because they provide a good way to signpost more inferred content. This finding supports the argument that using an uncontrolled vocabulary allows for a greater representation of marginalized identities. Visibility management is also an important factor to consider in regards to the subject metadata applied to books with queer subtext. Specifically, it can be subversive to minimize the visibility of books with queer subtext by not labeling them as such to extend their reach. When books and other forms of media contain queer subtext, consumers might not want others to know that this subtext is present – for example, if they are trying to privately explore their identity. Further, if books with queer subtext are not explicitly labeled as such, it can lead to readers being exposed to this content who otherwise would have not selected it, which can potentially lead to increased understanding and acceptance.

6.1 Limitations and future research

We acknowledge that our focus on picture books read during drag storytimes differs from the inclusion criteria of studies like Adler (2009), Bates and Rowley (2011), Williams (2017) and Johnson and Forsythe (2019). We believe that our approach uniquely captures the spectrum of diversity narratives present in children's picture books through its cultural and community context. This method embraces a wide array of themes, from books overtly focused on various facets of human difference to subtler explorations of difference and self-identity.

However, the specific books included in our sample may not be representative of children's picture books within the specific categories examined here more generally. For instance, as previously reported elsewhere (Barriage et al. 2024), nearly all books depicting a character with a DDDCI included characters wearing eyeglasses, with fewer books depicting characters with other types of DDDCI, such as characters who are blind or who use mobility aids. This may not be an accurate reflection of the larger body of children's picture books depicting characters with DDDCI; thus, our findings may not be generalizable to children's picture books more broadly.

Additionally, our analysis does not take into account the extent to which characters of various identities were integral to the books' storylines, aside from distinguishing between lead and non-lead characters. Diverse characters may be incorporated into picture books via what Izienicki (2022, 1100) has termed "backgrounding," with such characters included in the background of the story's main action (consider, for example, a Black character or a character who uses a cane depicted in a crowd scene on a single page). The inclusion of such characters may not warrant the application of relevant subject metadata in LCC or in LibraryThing. Future research examining identity-related subject metadata in children's picture books should include an analysis of the centrality of characters of various identities to each book's storyline and the influence this may have on the assigned subject metadata.

As previously noted, little work has examined subject metadata within the context of children's materials. In addition to diverse and fluffy subject metadata, nearly all records in both LibraryThing and LCC contained other metadata, or metadata that was unrelated to the diverse characters and/or themes present. This indicates that diversity is only one element considered sufficiently pertinent by both librarians and the general public when describing children's materials. Future research should continue to explore the application of subject metadata to children's materials by both information professionals and the general public, as well as the use of such metadata when retrieving children's materials for use in professional and personal contexts. Such

research could yield important insights that may influence the accessibility of diversity-related children's materials as well as children's materials more broadly.

7.0 Conclusion

Our comparative analysis demonstrates that surrogate records for books read during drag storytimes within the LCC and the LibraryThing platform contain subject metadata that is explicitly diverse, reflecting the various identities and themes they contain, as well as fluffy subject metadata, representing diverse content in more ambiguous or implicit terms. The prevalence of explicit and fluffy subject metadata is influenced by the type of book being described, as well as the type of catalog used. Librarians and drag performers selecting books to read during drag storytimes may want to tailor their searches accordingly in order to retrieve a range of potentially suitable options.

It is important to emphasize that we are not solely advocating for increased use of explicit subject metadata in the records of children's material to combat bibliographic invisibility. The application of fluffy terms may ultimately make diverse materials findable by embracing the methodological flexibility of uncontrolled vocabularies without supporting the use of prejudicial or inappropriate terminology. As with drag storytimes themselves, increasing the visibility of books with diverse content by applying explicit subject metadata may inadvertently make such material hypervisible (Kitzie et al. 2022), subjecting them to heightened scrutiny and negative attention. Although members of marginalized groups may be better able to find materials that reflect their identities, it also results in materials becoming findable by those who protest their inclusion in library collections. Individuals and groups who protest diverse material and programs such as drag storytimes may be able to find diverse materials more readily and call for their removal from the collection. This concern is not trivial, given the recent spate of book challenges across the United States that have primarily focused on books with LGBTQIA+ and/or BIPOC characters (Alter and Harris 2022; Stroshane, 2022). Catalogers should balance the desire to make diverse children's material more visible with the potential that doing so may make these materials hypervisible.

Endnotes

1. Here, in line with Williams (2017), we use the term 'fluffy' to indicate subject metadata that is "lacking in meaning or substance" (Merriam-Webster, n. d.).
2. Each of these books has also been identified as having the potential to be read as queer by other scholars and/or conservative groups that challenge the presence of LGBTQIA+ books in schools (Brand and Maasch 2017; Sullivan and Urraro 2017; Wargo and Coleman 2021; Stroshane 2022).

References

Adler, Melissa. 2009. "Transcending Library Catalogs: A Comparative Study of Controlled Terms in Library of Congress Subject Headings and User-Generated Tags in LibraryThing for Transgender Books". *Journal of Web Librarianship* 3, no. 4: 309-31. https://doi.org/10.1080/19322900903341099

Adukia, Anjali, Alex Eble, Emileigh Harrison, Hakizumwami Birali Runesha and Teodora Szasz. 2023. "What We Teach About Race and Gender: Representation in Images and Text of Children's Books". *The Quarterly Journal of Economics* 138, no. 4: 2225–85. https://doi.org/10.1093/qje/qjad028

Alter, Alexandra and Elizabeth A. Harris. 2022. "Attempts to Ban Books are Accelerating and Becoming More Divisive". *The New York Times*, September 16, 2022. https://www.nytimes.com/2022/09/16/books/book-bans.html

Association for Library Service to Children. 2020. "Competencies for Librarians Serving Children in Libraries". https://www.ala.org/alsc/edcareeers/alsccorecomps

Baldacchino, Christine. 2014. *Morris Micklewhite and the Tangerine Dress.* Illustrated by Isabelle Malenfant. Toronto: Groundwood Books.

Barriage, Sarah, Sammie Betler, Rebekah Lawler, Valerie Byrd Fort, Juniper Thorne, Vanessa Kitzie and Shannon M. Oltmann. 2024. "A Content Analysis of Picture Books Read During Drag Storytimes in Public Libraries." *Public Library Quarterly*. Published ahead of print. https://doi.org/10.1080/01616846.2024.2358609

Barriage, Sarah, Vanessa Kitzie, Diana Floegel, and Shannon M. Oltmann. 2021. "Drag Queen Storytimes: Public Library Staff Perceptions and Experiences". *Children and Libraries* 19, no. 2: 14-22. https://doi.org/10.5860/cal.19.2.14

Bates, Jo and Jennifer Rowley. 2011. "Social Reproduction and Exclusion in Subject Indexing: A Comparison of Public Library OPACs and LibraryThing Folksonomy". *Journal of Documentation* 67, no. 3: 431-48. https://doi.org/10.1108/00220411111124532

Bishop, Rudine Sims. 1990. "Mirrors, Windows, and Sliding Glass Doors". *Perspectives* 6, no. 3: ix–xi.

Bowker, Geoffrey C. and Susan Leigh Star. 2000. "Invisible Mediators of Action: Classification and the Ubiquity of Standards". *Mind, Culture, and Activity* 7, no. 1-2: 147-63. https://doi.org/10.1080/10749039.2000.9677652

Brand, Susan Trostle and Susan L. Maasch. (2017). "Updating Classroom Libraries and Cross-Curricular Activities: Celebrating Gender Identity and Diversity Through

LGBTQ Books". *Childhood Education* 93, no. 5: 430-9. https://doi.org/10.1080/00094056.2017.1367240

Bratt, Jessica Anne. 2022. *Let's Talk About Race in Storytimes*. Chicago: ALA Editions.

Brown, Christia Spears. 2021. *Unraveling Bias: How Prejudice Has Shaped Children for Generations and Why It's Time to Break the Cycle*. Dallas: BenBella Books.

Cahill, Maria, Erin Ingram and Soohyung Joo, S. 2021. "Storytime Programs as Mirrors, Windows, and Sliding Glass Doors? Addressing Children's Needs Through Diverse Book Selection". *The Library Quarterly* 91, no. 3*:* 269-284. https://doi.org/10.1086/714317

Cifor, Marika and K. J. Rawson. 2023. "Mediating Queer and Trans Pasts: The Homosaurus as Queer Information Activism". *Information, Communication & Society* 26, no. 11: 2168-85. https://doi.org/10.1080/1369118X.2022.2072753

Condren, Chelsea. 2018. "Far From a Drag: How One Library Embraced Drag Queen Story Hour". *Children & Libraries* 16, no. 1: 21–22. https://doi.org/10.5860/cal.16.1.21

Dobreski, Brian, Karen Snow and Heather Moulaison-Sandy. 2022. "On Overlap and Otherness: A Comparison of Three Vocabularies' Approaches to LGBTQ+ Identity". *Cataloging & Classification Quarterly* 60, no. 6-7: 490-513. https://doi.org/10.1080/01639374.2022.2090040

Dorsey, Zachary A. 2020. "When Fierceness and Kindness Collide: The Dramaturgy of a Drag Storytime". *Review: The Journal of Dramaturgy* 26, no. 1: 3-11. https://soundideas.pugetsound.edu/lmdareview/55/

Drabinski, Emily. 2013. "Queering the Catalog: Queer Theory and the Politics of Correction". *The Library Quarterly* 83, no. 2: 94-111. https://doi.org/10.1086/669547

Evans, Shane W. 2016. *We March*. New York: Square Fish.

Friedman, Batya, Peter H. Khan and Alan Borning. (2013). "Value Sensitive Design and Information Systems". In *Early Engagement and New Technologies: Opening Up The Laboratory,* edited by Neelke Doorn, Daan Schuurbiers, Ibo van de Pool and Michael E. Gorman, 55-95. Dordrecht: Springer.

Gough, Cal and Ellen Greenblatt. 1992. "Services to Gay and Lesbian Patrons: Examining the Myths". *Library Journal* 177, no. 1: 59-63.

Hauser, Elliot and Joseph T. Tennis. 2019. "Episemantics: Aboutness as Aroundness". *Knowledge Organization* 46, no. 8: 590-95. https://doi.org/10.5771/0943-7444-2019-8-590

Herthel, Jessica and Jazz Jennings. 2014. *I Am Jazz*. Illustrated by Shelagh McNicholas. New York: Dial Books.

Hoffman, Gretchen L. 2019. *Organizing Library Collections: Theory and Practice*. Lanham: Rowman & Littlefield.

Holley, Ralph M. and Daniel N. Joudrey. 2021. "Aboutness and Conceptual Analysis: A Review". *Cataloging & Classification Quarterly* 59, no. 2-3: 159-85. https://doi.org/10.1080/01639374.2020.1856992

Izienicki, Hubert. 2022. "The Big Picture: Representation of LGBTQ Characters and Themes in Picture Books Available in the United States 1972-2018". *Sexualities* 27, no. 4: 1091-116. https://doi.org/10.1177/13634607221144627

Johnson, Mackenzie and Carlie Forsythe. 2019. "Disability and Accessibility Language in Subject Headings and Social Tags". *Catalogue & Index* 197: 16–26. https://ir.lib.uwo.ca/fimspub/339/

Joudrey, Daniel N and Arlene G. Taylor. 2018. *The Organization of Information*. 4th ed. Santa Barbara, CA: ABC-CLIO.

Keenan, Harper and Lil Miss Hot Mess. 2020. "Drag Pedagogy: The Playful Practice of Queer Imagination in Early Childhood". *Curriculum Inquiry* 50, no. 5: 440-61. https://doi.org/10.1080/03626784.2020.1864621

Kitzie, Vanessa, Diana Floegel, Sarah Barriage, and Shannon M. Oltmann. 2022. "How Visibility, Hypervisibility, and Invisibility Shape Library Staff and Drag Performer Perceptions of and Experiences with Drag Storytimes in Public Libraries." *The Library Quarterly* 92: 215-240. https://doi.org/10.1086/719915

Larrick, Nancy. 1965. "The All-White World of Children's Books". *Saturday Review*, September 11, 1965. https://kgrice3.wixsite.com/lcyadiversity/all-white-world-1965

Loney, Andrea J. 2021. *Bunnybear*. Illustrated by Carmen Saldaña. Chicago: Albert Whitman & Company.

McClary, Carrie and Vivian Howard. 2007. "From 'Homosexuality' to 'Transvestites': An Analysis of Subject Headings Assigned to Works of GLBT Fiction in Canadian Public Libraries". *The Canadian Journal of Information and Library Science/La Revue canadienne des sciences de l'information et de bibliothéconomie* 31, no. 2: 149-62.

Mendell, Patricia and Patricia Sarles. 2010. "'Where Did I Really Come From?' Assisted Reproductive Technology in Self-Published Children's Picture Books". *Children and Libraries* 8, no. 2: 18-29. https://doi.org/10.5860/cal.8n2

Merriam-Webster. n. d. "Fluffy". https://www.merriam-webster.com/dictionary/fluffy

Montague, Rae-Anne and Joyce Latham. 2019. "Queer Reflections: New Views From Library Drag Storytimes". *IFLA WLIC*. http://library.ifla.org/2585/1/191-montague-en.pdf

Moulaison Sandy, Heather and Jenny Bossaller. 2017. "Providing Cognitively Just Subject Access to Indigenous Knowledge Through Knowledge Organization Systems". *Cataloging & Classification Quarterly* 55, no. 3: 129-52. https://doi.org/10.1080/01639374.2017.1281858

Naidoo, Jamie Campbell. 2018. "A Rainbow of Creativity: Exploring Drag Queen Storytimes and Gender Creative Programming in Public Libraries". *Children and Libraries* 16, no. 4: 12-22. https://doi.org/10.5860/cal.16.4.12

Nel, Philip. 2017. *Was the Cat in the Hat Black? The Hidden Racism of Children's Literature, and the Need for Diverse Books*. Oxford: Oxford University Press.

Nissenbaum, Helen. 2001. "How Computer Systems Embody Values". *Computer* 34, no. 3: 120-19. https://doi.org/10.1109/2.910905

Olson, Hope A. 2001. "The Power to Name: Representation in Library Catalogs". *Signs: Journal of Women in Culture and Society* 26, no. 3: 639-68. https://doi.org/10.1086/495624

Oltmann, Shannon M., Vanessa Kitzie, and Sarah Barriage. 2023. "'For Me, It Is an Intellectual Freedom Issue': Drag Storytimes, Neutrality, and ALA Core Values." *Journal of Librarianship & Information Science* 55: 734-743. https://doi.org/10.1177/09610006221100853

Palacio, R. J. 2017. *We're All Wonders*. New York: Knopf Books.

Radis, Brie, Katharine Wenocur, Jeffrey Jin and Colleen Keeler. 2022. "A Rainbow For Reading: A Mixed-Methods Exploratory Study On Drag Queen Reading Programs". *Journal of Creativity in Mental Health* 17, no. 3: 332–49. https://doi.org/10.1080/15401383.2021.1892557

Rolla, Peter J. 2009. "User Tags versus Subject Headings: Can User-Supplied Data Improve Subject Access to Library Collections?". *Library Resources & Technical Services* 53, no. 3: 174–84. https://doi.org/10.5860/lrts.53n3.174

Sipe, Lawrence R. (2008). *Storytime: Young Children's Literary Understanding in the Classroom*. New York: Teachers College Press.

Snow, Karen and Anthony W. Dunbar. 2022. "Advancing the Relationship Between Critical Cataloging and Critical Race Theory". *Cataloging & Classification Quarterly* 60, no. 6-7: 646-74. https://doi.org/10.1080/01639374.2022.2089936

Staino, Rocco. 2017. "Storytime Gets Fabulous". *School Library Journal* 63, no. 7: 14.

Stroshane, Eric. 2022. "Censorship Dateline". *Journal of Intellectual Freedom & Privacy* 7, no. 2: 45-72. https://doi.org/10.5860/jifp.v7i2.7968

Strottman, Theresa A. 2007. "Some of Our Fifty are Missing: Library of Congress Subject Headings for Southwestern Culture and History". *Cataloging & Classification Quarterly* 45, no. 2: 41-64. https://doi.org/10.1300/J104v45n02_04

Sullivan, Ashley Lauren and Laurie Lynne Urraro. 2017. "Missing Persons' Report! Where Are the Transgender Characters in Children's Picture Books?". *Occasional Paper Series*, no. 37. https://doi.org/10.58295/2375-3668.1095

Wagner, T. L. 2022. "'Describing Without Identifying': The Phenomenological Role of Gender in Cataloging Practices". PhD diss., University of South Carolina.

Wargo, Jon M. and James Joshua Coleman. 2021. "Speculating the Queer (In)Human: A Critical, Reparative Reading of Contemporary LGBTQ+ Picturebooks". *Journal of Children's Literature* 47, no. 1: 84–96.

Watson, Brian M. and Beck Schaefer. 2023. "Handicapped Has Been Cancelled: The Terminology and Logics of Disability in Cultural Heritage Institutions". *First Monday* 28, no. 1-2. https://doi.org/10.5210/fm.v28i1.12898

Westwater, Jason J., Elizabeth A. Riley and Gregory M. Peterson. 2019. "What About the Family in Youth Gender Diversity? A Literature Review". *The International Journal of Transgenderism* 20, no. 4: 351–70. https://doi.org/10.1080/15532739.2019.1652130

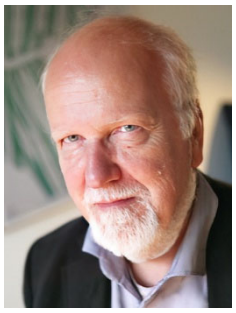
Williams, Sarah C. 2017. "Subject Headings Friend or Foe?: Cataloging Diverse YA Books in Public Libraries". Master's thesis, University of North Carolina at Chapel Hill.

Winner, Langdon. 1980. "Do Artifacts Have Politics?". *Daedalus* 109, no. 1: 121–36. https://www.jstor.org/stable/20024652

Bibliography (Field of Study)

Birger Hjørland

University of Copenhagen, Faculty of Humanities, Department of Communication,
South Campus, Building 14, 3rd. Floor, Karen Blixens Plads 8, 2300 Copenhagen S, Denmark
birger.hjorland@hum.ku.dk



Birger Hjørland holds an MA in psychology and PhD in library and information science. Since March 2020, he has been Professor Emeritus at the Department of Communication, University of Copenhagen and formerly professor in knowledge organization at the Royal School of Library and Information Science/Department of Information Studies (2001-2020) and the University College in Borås (2000-2001). He is a member of the editorial boards of Knowledge Organization, Journal of the Association for Information Science and Technology and Journal of Documentation. His h-index on 2024-12-16 is 52 in Google Scholar and 36 in Web of Science.

Hjørland, Birger. 2024. "Bibliography (Field of Study)". *Knowledge Organization* 51, no. 8: 700-711. 55 references. DOI:10.5771/0943-7444-2024-8-700.

Abstract: This article presents bibliography as a field of study. It consists of several traditions, of which enumerative (or "systematic") bibliography is considered most important in relation to information science, but at the same time tends to be rejected as a scientific or scholarly field by other bibliographical traditions. It is about making, using and evaluating bibliographies, which list all kinds of publications. Analytic, descriptive and textual bibliography are other subfields, which are important for establishing the identity of a given document (is the Hamlet that scholar A is reading the same Hamlet that scholar B is reading?) and for providing critical editions of important works. Historical bibliography (with the sociology of text and book history) is yet another subfield, a very broad one that lacks coherence, but which provides important perspectives on the functions of different kinds of publications. In the UNISIST model, bibliographies are considered secondary kinds of publications (based on primary literature and a prerequisite for tertiary literature). From the Library of Alexandria (c. 285- BC) to Google (and Google Scholar) it has been a utopian dream to establish universal bibliographical control, to make all publications relevant for those needing any special set of them. To optimize visibility and retrievability of documents is an important task for information science, related to the goal of bibliographical control, and to literature- and information searching in bibliographic and full-text databases. A theoretical view on (enumerative) bibliography was suggested by Margaret Egan and Jesse Shera in 1952, in which a new field called "social epistemology" was seen as a "parent" discipline for the study of bibliography. This view is critically examined in this article, and it is suggested that Shera's 1951 characterization of social epistemology represents a better foundation for bibliography.

Received: 21 October 2024; Accepted 29 November 2024.

Keywords: bibliography; information science; social epistemology; bibliographical control.

† Derived from the article titled "Bibliography (filed of study)" in the *ISKO Encyclopedia of Knowledge Organization*, Version 1.0, published 2024-10-15.

1.0 Introduction

The term "bibliography" is used about a kind of document (which is characterized by focusing on providing bibliographical references to published documents), as well as about a field of study (or fields of study). This article focuses on the last sense, while an article about the first sense is planned as an independent article.

Foot (2006) found that bibliographical studies originated in nineteenth century with an emphasis on "enumerative bibliography," which means the norms and processes of making lists of publications and the typologization and evaluation of such lists. From the end of the nineteenth century the field widened out to include historical bibliography and the study of books as material objects. In the mid-twentieth century this wider approach narrowed down, as a consequence of much emphasis being placed on descriptive, analytical, critical and textual bibliography, but again widened out under the influence of French book historians. These different fields (or subfields, approaches or traditions) of bibliography are briefly presented in Section 2 of the present article.

The relations between bibliography on the one hand, and on the other hand information science (here considered synonymous with library and information science, LIS[1]) with knowledge organization (KO) are important[2]. What

today is called "information science" had, according to Kline (2004, 19), "bibliography" as one of its former names:

> Called bibliography, documentation, and scientific information during the first five decades of the twentieth century, the field became known as information science in the early 1960s[3].

One of the most important indicators of the relationship between documentation and information science is the change in name of the *American Documentation Institute* (founded in 1937) in 1968 to the *American Society for Information Science* (today the *Association for Information Science & Technology*, ASIS&T)[4]. The term LIS arose in connection with the inclusion of information science by library schools from 1964 until almost all schools had changed to LIS by the end of the 1990s (cf., Galvin 1977).

"Documentation" is thus a former name of information science, and it is closely related to bibliography[5]. The founder of the documentation movement, Paul Otlet, founded the *Institut International de Bibliographie* (IIB) in 1895, and wrote articles about bibliography as a science (Otlet 1990a; 1990b), which understood bibliography to be about publications in general, not just about books. Otlet wrote (1990b, 86):

> The Science of bibliography can be defined as that science, whose object of study is all questions common to different kinds of documents: production, physical manufacture, distribution, inventory, statistics, preservation, and use of bibliographic documents; that is to say, everything which deals with editing, printing, publishing, bookselling, bibliography, and library economy. The scope of this science extends to all written or illustrated documents which are similar in nature to books: printed or manuscript literary works, books, brochures, journal articles, news reports, published or manuscript archives, maps, plans, charts, schemas, ideograms, diagrams, original or reproductions of drawings, and photographs of real objects.

Otlet created, with Henri La Fontaine, the *Universal Decimal Classificatio*n (UDC), which illustrates its close connection to the field of knowledge organization[6].

There are indications that bibliography became less influential in relation to LIS when the field changed name to information science. Bibliography was considered a core element in schools of library science, but gradually the study of information behavior came to play a larger role at the expense of courses in bibliography. There were even some voices claiming "the bibliographical paradigm" to be obsolete (e.g., Henri and Hay 1994), a view which was counter argued by Hjørland (2007).

Michael Buckland raised the question (referred by de Fremery 2024, 1): "What might be gained by reinvigorating bibliography?". This question indicates that bibliography has lost influence in information science, and that this loss may have been harmful. This question was raised by de Fremery (2024), by Hjørland (2024a), and is also central in the present article.

2.0 Different subfields of bibliography and their relations to LIS

2.1 Enumerative/systematic bibliography

Hjørland (2024a) found that the most important distinction is between enumerative bibliography (also called systematic bibliography and reference bibliography) on the one hand, and all the other bibliographical traditions on the other hand. Compared to other parts of bibliographical studies, enumerative bibliography seems to lack proper theory as well as recognized bibliographers[7] (one of the leading persons, Theodore Besterman, 1904– 1976, was more a compiler of bibliographies than he was a theorist; another, D.W. Krummel displayed a research attitude and important insights [e.g., in his 1984 book[8]], but overall his contributions seem insufficient to form the theoretical basis for enumerative bibliography as a research field). Leading bibliographers, such as W. W. Greg and Donald Francis McKenzie, considered bibliography a science, but did not include enumerative bibliography in this science, as they recognized its utility as a separate activity[9] (it is well-known, that a list of bibliographical references in itself cannot be accepted as a research contribution in academia). Hjørland (2024a) argued however, that enumerative bibliography is the part of bibliography, which is most important for information science, and that information science with knowledge organization has provided the most important theoretical works which are relevant for this field. In other words enumerative bibliography is more connected with LIS than with other traditions of bibliography. The theoretical contributions from LIS include classification, indexing, and metadata assignment (or resource description) for documents in bibliographical databases, search strategies, recall and precision as evaluation criteria, bibliographical control, the use of references in literature searching, and much more. The term systematic bibliography indicates that candidates for inclusion in the bibliography are based on systematic criteria (e.g., lists of journals being indexed). Systematic bibliographies are necessary for systematic searches, and thereby also for systematic literature reviews, which illustrates its close connection to LIS.

There are various ways of classifying enumerative bibliographies, for example, national bibliographies, subject bibliographies, author bibliographies, and bibliographies lim-

ited to certain kinds of documents such as dissertations, journals, journal articles, books, and maps. As already said, bibliographies as kinds of documents fall outside the scope of the present article, but are mentioned here because the study of national bibliography, subject bibliography, etc. are parts of the field of enumerative bibliography.

Hjørland (2024a) concluded:

> A domain in which systematic literature searching and thereby enumerative bibliographies and bibliographical control is taken most seriously is evidence-based medicine, where knowledge of the most important findings is of utmost importance. Much research is carried out about databases coverage of relevant findings, about retrieval strategies etc. Such research, about the bibliographical coverage and findability of documents relative to a research paradigm is a core issue for a theoretical research in enumerative bibliography as well as in information science.

The relation between enumerative bibliography and literature searching is further developed in Section 5.

2.2 Analytical, descriptive and textual bibliography

2.2.1 Analytical bibliography

Analytical bibliography studies the processes of making books, especially the material modes of production. One of its purposes is to show how the processes of material production affect the nature and state of the text preserved in the book. A main representative is Philip Gaskell (1974) *A New Introduction to Bibliography*. This book covers hand printed and machine printed books through the ages.

2.2.2 Descriptive bibliography

Descriptive bibliography emphasizes details about page layout, typefaces, bindings, and other elements that help identify a book's edition[10]. It draws on analytical bibliography. Probably the main representative of this field is Fredson Bowers' (1949) book *Principles of bibliographical description*. As exemplified by de Fremery (2024, 184), the descriptions "created by Bowers and Greg made it possible to know that the Hamlet discussed by scholar A is the same Hamlet discussed by scholar B".

Three comments should be made. First, it is clear that not every writer can be carefully studied in the way that descriptive bibliography suggests. For average writers, the task of making precise identifications of versions of their works, must be done by themselves in cooperation with publishers and editors. As with the example with two scholars discussing Hamlet, ordinary researchers need to know if the text they read and cite is the same as another author has cited. It is a sign of bad scholarship if the edition or version of a document is not made explicit and precise, for example, when articles are reproduced in edited books.

Second, in relation to electronic documents, two electronic copies of the same text can be considered entirely the same because each bit is checked in the copying process, and the probability that two copies are not exactly similar is extremely small. However, different instantiations[11] or versions often have been published, and even the same file does not interact in the same way with different versions of the software used to display it. These facts are issues to consider in relation to digital documents related to the issues of traditional descriptive bibliography. Gants (2010) presents principles for description of electronic publications analogue to the principles developed by Bowers (1949).

Third, "descriptive bibliography" is not about the description of contents of publications, such as done, for example, by "abstracting journals" (e.g., Chemical Abstracts, MEDLINE, and PsycINFO). Neither is this field about descriptive cataloging of documents (see Tanselle 1977; Yee 2007) concerning the relations between descriptive bibliography and library cataloging). The ways of referencing in academic books and journals, such as ISO 690, "Harvard system", "APA-style, "Chicago style", "MLA Style", and "Vancouver system" or in electronic referencing systems such as "EndNote," "Zotero," "Reference Manager," "RefWorks" and "ProCite" have also failed to attract the interest of bibliographers as well as information scientists.

2.2.3 Textual bibliography

Textual bibliography is also called "critical bibliography". It was defined by Reimer (2015):

> Textual bibliography attempts to establish the "state" of a text, especially in terms of the various versions that are extant, and analyzing who (author, editor, compositor, printer, etc.) was responsible for particular variants. Textual bibliography is obviously part of the process of preparing a scholarly edition of a text, though its significance is certainly not limited to editors.

Textual bibliography is used to produce "critical editions," which are attempts to construct a text of a work using all the available evidence. Prominent examples are studies of the Bible, of Shakespeare and other "Great Books" in different cultures. Such bibliographical studies have often been extremely important for subsequent researchers. Among the influential works in this field is G. Thomas Tanselle (1990), *Textual Criticism and Scholarly Editing*.

2.2.4 Considered together

Considered together, analytical, descriptive and textual bibliography do not have the same importance for LIS as enumerative bibliography has. However, libraries with collections of old and rare books have an interest in such studies in order to identify different versions of the books, and in order to support scholarly research based on them[12]. Therefore, these areas of bibliography are important in relation to the management of some kinds of collections. In addition, principles and concepts developed in these areas may have importance for other areas of bibliography as well as for LIS (Tanselle 1977 advocated for a cooperation between bibliographers and cataloguers).

2.3 Historical bibliography / sociology of texts/ book history

Historical bibliography examines the history of the book as a cultural artifact. It explores how books and other documents have influenced and been influenced by historical contexts, including how they reflect and affect social, cultural, and intellectual movements. It includes the evolution of book production and dissemination over time as well as the history of reading practices. McKenzie (1999) focuses on how texts are intended by their authors as well as on how they are received and interpreted by readers. He advocates for a view of bibliography that not only acknowledges the technical and physical aspects of texts but also fully embraces their social and cultural dimensions. This is a wide field, which has been difficult to define. Krummel (2017, 479) found that if there are differences between historical bibliography and the new fields of study called "print culture" and "book history," they are subtle and often irrelevant.

Information science also has interest in these kinds of studies, especially as this relates to scientific and scholarly communication and the roles different kinds of documents play in domains, between domains and in relation between science and the broader society. Although this part of bibliographical studies is broad, and in lack of coherence, it is important in order to obtain a deeper understanding of the fundamental issues in information science.

3.0 Bibliographies according to the UNISIST model

Bibliography is one among many types of documents. The UNISIST model, originally published in 1971, later revised and updated by Fjordback Søndergaard et al. (2003), offers an important sociologically-oriented perspective on the activities of scholarly communication. It seeks to draw attention to information communication between knowledge producer and knowledge user, as a system consisting of diverse organizational and documentary units each contributing to the division of labor in scholarly communication. It provided a model of scientific and technical information services and document types, in two dimensions, of which the most important is based on three levels: (1) primary sources and services, such as books and journals mediated by publishers (2) secondary sources and services, such as catalogs, bibliographies, and abstracting and indexing journals provided by libraries, clearing houses and producers of bibliographic databases (3) tertiary sources and services, such as systematic reviews and other forms of syntheses of the primary literature.

A model such as the one provided by revised UNISIST Fjordback Søndergaard et al. (2003) put bibliographies in the context of different institutions, services and document types, and thereby raises important questions such as, which documents from which disciplines are included in different bibliographies? What are the relative roles of bibliographies for users in obtaining the documents they use? The answers are domain specific, where, for example, evidence based medicine (EBM) is a field in which bibliographies are extremely important, while, for example, many fields in the social sciences do not rely on bibliographies in the same way.

4.0 Bibliographical control/bibliographies as infrastructures

For some purposes, bibliographical control is important. In general it is considered important in academia that knowledge production is based on knowledge about what has already been written or documented about a certain topic. Throughout history there has been a utopian dream of organizing all publications (or all "information") and make it possible to identify all relevant documents. Hjørland (2023) discussed the following historical examples of attempts of fulfilling this dream:

- The Library of Alexandria (c. 285- BC)
- Gessner's Bibliotheca Universalis (1545)
- The abstract journal and subject bibliographical databases (1790-)
- The Universal Bibliographical Repertory, RBU (1895-)
- Union catalogs (1930s-) with WorldCat (1971)
- IFLA's and UNESCO's program on universal bibliographical control (1970s-)
- The World Wide Web (1989-) and Google

Each of these attempts can also be said to represent important bibliographical infrastructures, and an important task for information science with knowledge organization is to study the effectiveness and efficiency of such infrastructures, to communicate about them to potential users, and contribute

to their improvements. These infrastructures are seldom examined from a holistic perspective, but the so-called "Sputnik crisis" in 1957 made important attempts in the USA to improve its systems of scientific and technological communication, and produced influential reports, including the Weinberg Report (President's Science Advisory Committee 1963) and the SATCOM report (Committee on scientific and technical communication 1969) (see also Rayward 2024 about the overall development of attempts to create international bibliographical control or "information order" since World War II). It seems important that the field of information science maintains a holistic focus on bibliographical infrastructures and their effectiveness and efficiency.

5.0 Literature- and information searching

Libraries have for more than one? hundred years had important tasks helping users finding the documents they need for their activities. They have done so by designing their catalogues to serve this task, they have provided open collections of reference works with bibliographies and they have made bibliographic guides (now including "Libguides") and provided courses in literature searching for students in different disciplines. Library schools were active in this field, for example teaching (enumerative) bibliography as a subject, introducing kinds of bibliographies as well as important examples of concrete bibliographies. Professional organizations, such as the American Library Association (ALA) contributed, for example with publications such as Webb et al. (1986) *Sources of Information in the Social Sciences. A Guide to the literature*. In short, bibliography was very important in this period. Before the online period, this was typically called "literature searching", in the online period typically changed to "information searching" or "information retrieval," although mostly the same bibliographical databases were used (e.g., *Psychological Abstracts* founded in 1927, but changed to electronic format and called *PsycINFO* from 1967).

When the online industry developed from about 1963, their services were partly offered by research libraries, and partly by so-called "information centers". The formerly influential database host DIALOG organized its databases in five main groups (Niro et al. 2008, 5):

- Bibliographic databases. Each record in a bibliographic database is a reference or citation (many also include a summary or an abstract) to a publication, magazine or journal article, news story, patent, conference paper, etc.
- Numeric databases. Each record in a numeric database is a table of statistical data, often with text added.
- Directory databases. Each record gives factual information about companies, organizations, products, chemical compounds, etc.
- Complete text databases [fulltext databases] Each record includes the complete text of magazine [or journal] articles, newswire stories, patents, etc. [Books, encyclopedias]
- In addition, some databases contain a mix of several different kinds of data, such as bibliographic and fulltext records.

Among these groups, the bibliographic databases played by far the most important role in this period. In library schools (which from 1964 gradually changed their names to schools or departments of library and information science, LIS) the teaching of bibliography now typically included online searching (an influential textbook was Harter 1986), which clearly contributed to increase the status of LIS/information science.

Around 1990 began a new development when full-text databases became common (which was primarily due to falling costs of storing information on computer discs) and the development of the Internet and search engines. From that time full-text retrieval began to compete with databases that only contains bibliographical records. Bibliographical searching still is important, and, for example in EBM are "classical" databases like MEDLINE still considered core information sources (see Higgins and Green 2009 and updates). The core competencies of information specialists involve both bibliographical databases and fulltext-databases, and the basic principles are not much different. What *is* different, of course, is the tendency to let algorithms and artificial intelligence perform the searches. Hjørland (2015) argued however, that such systems, although very user-friendly, do not provide the necessary transparency and control over the search process for important purposes such as EBM. It is therefore no exaggeration to say that core competencies of information professionals to select, use and evaluate information systems are still closely tied to the concept of bibliography.

The research front in the study of bibliography is visible, for example, in EBM, where much research is carried out about databases coverage of relevant findings, about retrieval strategies etc. Such research, about the bibliographical coverage and findability of documents relative to a research paradigm, is a core issue for theoretical research in enumerative bibliography as well as in information science.

6.0 Theory of enumerative bibliography

Enumerative bibliography is by some researchers considered a part of bibliographical studies (e.g., Foot 2006), while other researchers have denied this connection (e.g., Greg 1930; Bowers 1949) a. It is generally considered theoretically weak compared to other bibliographical traditions. However, as stated in Section 2.1, it is closely related to in-

formation science. Here we shall consider its theory with the point of departure in Egan and Shera (1952) "Foundations of a Theory of Bibliography," which suggested (131-134) that "social epistemology," an envisioned new field of study should function as a "parent" discipline for [enumerative] bibliography. This paper is often cited as the first occurrence of the term "social epistemology," but as documented by Hjørland (2024b) the term was used by Shera (1951, 82) in the context of classification and in another sense. While the 1951 understanding is considered valuable, the understanding provided by Egan and Shera (1952) seems muddled. The analysis of why this is the case can shed light on the theory of enumerative bibliography.

Egan and Shera (1952) aimed to base the theory of bibliography on the needs of users, and wrote (135):

> The first basic need is for a complete analysis of the kinds of information, knowledge, and insights developed by all the contributory sciences or disciplines that are brought to bear upon each of the many focal points of human activity. Such analyses would answer, for example, such questions as: What information or knowledge is required when a business enterprise or commercial undertaking proposes to open a new market? What information or knowledge should be available when a legislative body is considering a new or revised tax law? What information or knowledge is essential to a chemical industry that is developing a new synthetic fiber? Such situations might be multiplied indefinitely; although each is unique, all probably fall into a finite number of discoverable types distinguished from one another by the possession of identifiable characteristics.

The authors recognized an important problem (135):

> That exploration of this kind cannot be done once and for all is obvious, for the situations themselves will change as the factors that condition them change. An illustration will make this clear. Let it be assumed, for example, that someone should suddenly establish beyond all doubt that the cause of poliomyelitis is not bacteriological but systemic and chemical. Such a development would be immediately reflected in a drastic alteration of the informational needs and requirements of those working to improve the diagnosis, treatment, cure, and prevention of this dread disease.

The example says "suddenly establish beyond all doubt," but why this reservation? Why not say: "the scientific theory that the cause of poliomyelitis is bacteriological is replaced by a new theory that it is systemic and chemical". The example therefore suggests that the information needed change when the theory in the domain of enquiry changes. This would require that bibliographic services are tailored to theories rather than to users, a view that corresponds to the domain analytic perspective (cf., Hjørland 2017). In a given field, for example, art studies or psychology, there may be competing theories and "paradigms", and thereby a need for competing bibliographical services, in psychology for example cognitive, psychoanalytic and sociocultural services. The providers of bibliographical services are therefore necessarily parts in the epistemological struggles in the domain (therefore, as pointed out by Hjørland (2024b), it was very unfortunate, that the epistemological understanding of social epistemology in Shera (1951) was replaced by a sociological understanding in Egan and Shera 1952).

Egan and Shera were right when they suggested social epistemology as a parent discipline for enumerative bibliography, but unfortunately they forgot the important epistemological perspective suggested by Shera (1951), which, according to Hjørland (2024b, 194) implied:

1. That a classification cannot be universal, serving all purposes for everybody that a classification cannot be permanent;
2. That classificationists build on their predecessors, making classifications developmental and dynamic;
3. That classifications are based on "the materials at hand", i.e., based on the knowledge and concepts of its time;
4. That classifications are designed to serve specific needs.

Hjørland (2024b, 194) continued:

> Shera's expression "the intellectual environment of its age" may be translated to the dominant worldview, paradigm, epistemology, or metatheory. Probably, we should not take "generation" too literally. It may well be that some classifications have a longer and some have a shorter lifetime and that not all fields of knowledge necessarily develop in a synchronized way. What is important is that the classifier (and the resulting classification) is influenced by views represented in a broader social, cultural, and domain-specific context. This is a clear social epistemological position that denies the possibility of constructing classifications based on the isolated individual's observation and cognition.

This view is opposed to Otlet's (1990b, 85) view, that considered that bibliography and documentation should communicate scientific information "in an analytic form from which any personal interpretation has been removed". Otlet's article even suggested that books would become obsolete, and the production of knowledge would change to publication of catalog cards and loose-leaf publications,

which should constantly be updated and replace older catalog cards and loose-leafs[13]. Such thoughts reflect a positivist philosophy, which is strongly opposed to Shera's social epistemology.

7.0 Conclusion

What, then, is bibliography? The conclusion here is that that this term does not relate to a single field of study, but to a family of fields as presented in Section 2. Some of these areas are highly interrelated, whereas enumerative bibliography has a weaker connection to these highly interrelated fields, but has a tighter connection to LIS. Therefore, we cannot expect a single definition of bibliography as a field of study to be found[14].

What, then, is enumerative bibliography (synonyms: systematic bibliography and reference bibliography)? Like the other fields labeled "bibliography" in Section 2, enumerative bibliography is about the study of documents (or "information resources"). The focus of enumerative bibliography is to facilitate information retrieval by listing, analyzing and describing documents in order to enable users to identify the documents needed for a specific task. The concept of "bibliographical control" is central, just as the synonym "systematic bibliography" as already said, indicates a relation to "systematic search" and "systematic review".

Enumerative bibliography, in some form or another, is an essential tool for science and scholarship. Hjørland (2012, 63) wrote (translated by the author):

> All good research – and all good, independent university assignments – begin and end in the literature of the discipline.

If a person P is educated in discipline X, this means that the research done in X is what makes P a professional person (as contrasted to an amateur). P may draw on other disciplines, but the point of departure is the field in which P is educated. P may find serious gaps or problems in the literature of X (in general or just in relation to P's topic), and therefore P may draw on other disciplines or on his own common sense. This implies, however, an attempt to improve X, and if P argues for this in the literature of X it may constitute a contribution to X. If nobody contributes to X, and if claims in X are not examined, then X will be sterile, obsolete, or simply useless. This implies that people educated in X, including P, cannot be a competent professional, although they may be unaware of this because they have not examined the basis of what they have been taught. This process also implies that a contribution to X is not just an isolated fragment, but is at the same time an argument about developing X in a certain direction. There are always plenty of different directions in which a given discipline can be developed, but if too many ways are followed at the same time, it means that the discipline is dominated by what Cronin (2002) and Hjørland (2013, 208) called "centrifugal forces" bringing about conceptual fragmentation with the risk of a dissolution of the field. Any discipline therefore needs a balance between centripetal and centrifugal tendencies[15] in its research fields.

It follows that in well-functioning disciplines or specialties, researchers are interested in following what their colleagues are doing, and the directions in which the field is developing. For this purposes, different kinds of bibliographic tools are important for current awareness as well as retrospective searches and citation networks. In order to support these activities in an optimal way, bibliographic tools must be based on knowledge of theoretical positions in the field, and a conscious priority. Billig (2013) suggested that in order to success in the social science, researchers must learn to write badly. Of course, this is not a goal that an ideal bibliographical services should aim at supporting, but Billig's view should not be considered a joke, but a serious criticism of the state of the art in these domains, which is also important to know about for information and KO professionals.

Hjørland (1992, 189) concluded: "Thus an analysis of a subject is itself, at its most profound, a part of the scientific process of knowledge gathering". This can be generalized to bibliography as a whole: The provision of bibliographical services and product, is, in the end, a matter of supporting the activities of researchers and the theoretical development of the discipline or specialty. If, for example, information specialists are unable to distinguish between bad and good writing in the way suggested by Billig, they are in a bad position to provide quality information services[16]. In short, a subject bibliography is a kind of map of a subject domain, and as is the case with other maps, it cannot be a neutral mapping[17]. This view corresponds to Shera's 1951 social epistemology, and is opposed to Otlet's positivism as described above.

Acknowledgements

Thanks to Daniel Martínez-Ávila, who served as the editor of this article, and to three anonymous referees, the comments, in two cases detailed comments, contributed to improving this article.

Endnotes

1. Information science and LIS are not always considered synonyms, and "information science" is an ambiguous term. Hjørland (2013, 223-224) wrote: "In 2002, two different international conferences about the foundations of information science took place. One was the Fourth Conference on Conceptions of Library and Information Science (CoLIS 4) in Seattle, USA, the other

was the International Conference on the Foundations of Information Science (FIS).26. Were these conferences discussing two different fields, each of which claimed to be an "information science", or were they two different scholarly meetings in the same field? Perhaps they are both forums for multidisciplinary approaches using different disciplinary outlets? Whether they represent one, two, or more kinds of information sciences can only be uncovered by theoretical analysis of the core assumptions expressed in the respective conferences and their proceedings. Inasmuch as FIS is founded on cybernetics and CoLIS is founded on something more related to social and epistemological studies of knowledge production and dissemination, different information sciences may well be at play".

2. One of the anonymous referees objected to the view that KO is considered a subfield of information science. The definition and delimitation of fields/terms such as KO, information science, library and information science (LIS), documentation, bibliography etc. is a complex task. I have formerly written about this in, for example, Hjørland (2013; 2016; 2018; 2018b) and I find that a further discussion of this falls outside the scope of this article. The referee indicated that (s)he considers KO to be a science in itself. I have some reservations about this view. As far as I know, "knowledge organization" (or "information organization") only exists as courses in the field of LIS/information science, and therefore must be considered a subfield of LIS (or information science). One cannot, in my opinion speak of a science (or discipline or field of knowledge) as a purely hypothetical, or possible, or logically given thing. ISKO is a community of KO, but it is not in itself enough to define a scholarly discipline (let alone given theoretical divergences within ISKO). Also, the term "science" seems excessive; field of study or discipline are more modest.
3. An anonymous reviewer wrote: "Kline's (2004) statement, which suggests that Bibliography is an old name for Documentation and that Documentation is an old name for Information Science, lacks argument and needs to be questioned". Well, I admit that the development is complex and not unambiguous. It is however outside the scope to provide a more detailed account (see e.g., Hjørland 2013). Kline's quote is made in order to state the relation between bibliography and LIS (with KO). Some other comments from the reviewer seem-not to indicate that (s)he disagrees that there is a close connection between these fields.
4. One of the anonymous referees wrote: "Basing oneself on changes in the names of institutions, such as the change of the Institut International de Bibliographie (IIB) to the International Federation for Documentation, and of the American Documentation Institute to the American Society for Information Science, does not theoretically support the idea that Bibliography became Documentation and that it would later become Information Science. This is merely a perspective that seeks to understand scientific fields as excessively linear in a causal relationship". I agree with the reviewer that it is important to distinguish between the institutional constitutions of fields of knowledge and their theoretical constitutions. I have in former writings, e.g., Hjørland (2000), objected to the institutionalization of the term "information" rather than "document"/ "documentation". However, each of us cannot operate with our individual languages, but has to consider the common languages as institutionalized in university departments, in journal names, in scientific societies etc. Therefore, I use names in institutions such as the *Institut International de Bibliographie* (IIB), the *International Federation for Documentation*, *the American Documentation Institute* and the *American Society for Information Science* as indications of overall trends. I see such name changes as neither linear nor causal, but just as representing a dominant but problematic trend or paradigm. I fully agree with the reviewer that the field of bibliography has a complex history, including the history of "informatika" in Russia. But when we use terms, we cannot each time produce a listing of different meanings, but have to suggest our own concepts in an argument with the dominant meanings in the context where we live.
5. Buckland (2021, x): "The resolution was to adopt document as the term of choice for any and all objects regarded as signifying, as evidence of anything. By extension, during the 1930s documentation came to replace bibliography as the term of choice in Otlet's circle and elsewhere. What began as an International Institute for Bibliography in 1895 became an International Institute for Documentation (IID) in 1931 and the International Federation for Documentation in 1937". Compare, however, Buckland (2021, xiv): "Paul Otlet was dedicated to an expansion of the interest in physical forms of documentation beyond printed documents to include administrative records, statistical data sets, heritage objects, specimens, and records of every kind".
6. An anonymous referee wrote: "The gloss that Otlet created ... the UDC ..." is quite a gloss. The Smiraglia and van den Heuvel piece (2013) provides more context. Otlet was quite concerned with KO as a basis for many things, right down to architecture (see the work of Wouter van Acker for example, start with his 2011 PhD dissertation).
7. Biagetti (2020) is an article suggesting the historical-bibliographical paradigm as a supplementary paradigm

in knowledge organization. Her article is mainly based on Alfredo Serrai's monumental work Storia della Bibliografia in 11 volumes (Serrai, 1988-2001). Biagetti's article suggests that Serrai might be the most important researcher in [enumerative] bibliography, as his research is based on a careful study of bibliographies and library catalogs over centuries, emphasizing the development of indexing theory (mainly based on rhetoric). Unfortunately this work is only available in Italian. Although it has received citations in the English-language literature (including a few in LIS/KO journals) Serrai is still not a well-known name in enumerative bibliography.

8. As pointed out by an anonymous reviewer, Krummel (1986) treated the whole field of enumerative bibliography in some detail, from the situation of the academic purposes of bibliographies to the enumeration of citation styles, annotation, and internal organization.
9. In the field of bibliography Cowley (1939, 6) wrote that "subject bibliography" in his opinion is not really bibliography, and therefore this term should not be used about these kinds of work, but should be replaced by, for example, "register", "catalogue", or "guide". He wrote: "[In this] kind of compilation [...] Minute description of the physical form of the material is therefore out of place, whereas criticism of its subject-matter is all-important". It is important to consider, however, that the kind of bibliography which gave rise to the documentation movement and thereby to information science and library and information science (LIS) with knowledge organization (KO) is the kind which Cowley dismisses.
10. An example of how descriptive bibliography typically describes the contents of books is given by Cowley (1939, 120): "Analysis of the Contents. The object of the contents paragraph is to provide a complete description of all the literary contents of the book or rather all the printed parts of it, including preliminaries, text, appendices, tables, etc., and to indicate their places within the physical framework. Hence every printed part is mentioned in the order in which it occurs and a reference is given to pages on which each part begins and ends. For this purpose every piece of printed matter is treated as of equal importance, even if it has no real significance for the literary value of the book, and every page must be accounted for, including blanks, if a complete description is being written. In subject bibliographies only the text and such subsidiary parts as modify the value of the text need to be mentioned".
11. Concerning the concept "instantiation," see Smiraglia (e.g., 2001; 2017).
12. An anonymous reviewer commented: "When the author states that descriptive, analytical and textual bibliographies do not have the same importance as enumerative bibliography for Library and Information Science, except in cases of libraries that have collections of rare and old books, he seems to neglect the fact that Library and Information Science courses also concern themselves with subjects corresponding to book publishing, book history, bibliographic publications, among others". Yes, this is correct, but still does not change my view that enumerative bibliography is closer related to LIS that it is to analytical and descriptive bibliography, and for LIS is the most important part of what goes under the name of bibliography.
13. This description of Otlet's philosophy represents an extreme example, which is not intended to characterize all of his opinions. The example is presented to illustrate contrasting philosophical positions in bibliography, not as an overall evaluation of Otlet's work, which should be the topic of an independent biographical article.
14. One of the anonymous reviewers suggested a definition of bibliography. He suggested:
    "– [1] the act of resource description even in library cataloging is derived from bibliography and follows a set of particulars known by various names, the most well-known being a "formulary;"
    – [2] the formulary (or rules) yield a description that matches precisely the details of specific artifacts (items) and of the works contained in them;
    – [3] the structuring of retrieval systems (files, bibliographies, catalogs, etc.) for works constitute an alphabetico-classed form of KO in which the class is named for a creating entity, divisions identify the works created by that entity, and subdivisions identify the specific items/artifacts that contain and transmit those works;
    – [4] the formularies themselves constitute a form a set-theoretic that is essentially the "science" of bibliography".

    Comments: (1) The relation between library cataloging and bibliography has, according to Tanselle (1977), been intrinsically complex. He regrets the lack of contact (and the sometimes unsympathetic, or even Hostile attitudes, toward each other's practices) and argues for a cooperation between the two fields. Therefore, I doubt we can say that library cataloging is derived from bibliography, although there probably has been some influence. (2) The cataloging rules are influenced by changing ideas, theories and ideals in the library community. There is no reason to believe that one best way of describing documents have been reached, rather, any set of rules always tends to serve some interests at the cost of other interests. It is one of the tasks of bibliography to examine the functionality of different descrip-

tions in relation to various goals and interests. (3) It is hard to understand the reviewer on this point. For me, what is constructed in catalogs and other kinds of bibliographical databases are bibliographical records. Such records provide different kinds of access points for information searching, including subject access points (see Hjørland and Kyllesbech Nielsen 2001). (4) The optimization of bibliographical records may be considered one task for the field of bibliography, but bibliography (but not cataloging) may be also about selecting the documents to be included in bibliographies, and about developing different kinds of bibliographical tools for different purposes and users.

15. Corresponding to what Becher and Trowler's 2001 discussed as "convergence" and "divergence" in research fields.
16. In national bibliographies formal criteria, not quality criteria, are used to define what should be included. In *Web of Science* the dominant selection criterion is the journal impact factor (JIF). Such examples seem to contradict the statement "if information specialists are unable to distinguish between bad and good writing, they are not able to provide quality information services". A defense for this thesis is that (1) national bibliographies may serve the book trade well, but be a poor tool for finding literature by researchers (2) That the use of JIFs as selection criterion implies a hypothesis that can be questioned. Probably this criterion is used because it is a cheap and easy way to manage the selection problem. Information specialists' knowledge about good and bad writing may be utilized in indirect ways, for example, by make methodological, epistemological and related characteristics visible for the searcher.
17. *The Times Comprehensive Atlas of the World* (2014, 42) wrote: "The power of maps. Maps are an extremely powerful form of geographic representation. Maps define territory – they tell of ownership and domination, they marshall spatial information. They can also subvert and propagate alternative worldviews. All maps serve an interest and work through two main forms of power. First, the external power of their creators, often governments and their agents, who control the content of maps both in terms of what is included and what is withheld, and thereby broadcast a particular viewpoint. Second, the internal power of maps themselves – the perception of maps as precise, objective and accurate representations of reality which convey an image of geographical order. Maps are still regarded by many people as dispassionate representations of the external world. However, this has been challenged in recent decades as their political and cultural connotations are revealed and become more widely understood".

References

Becher, Tony, and Paul R. Trowler. 2001. *Academic Tribes and Territories: Intellectual Enquiry and the Culture of Disciplines*. 2nd ed. Buckingham, UK: The Society for Research into Higher Education.

Biagetti, Maria Teresa. 2022. "The Indexical Bibliography: A Historic-Bibliographic Paradigm for Knowledge Organization". In *Knowledge Organization across Disciplines, Domains, Services and Technologies: Proceedings of the Seventeenth International ISKO Conference 6-8 July 2022 Aalborg, Denmark,* edited by Marianne Lykke, Tanja Svarre, David Haynes, Mette Skov, Martin Thellefsen and Daniel Martínez-Ávila. Advances in Knowledge Organization 18: Ergon: Baden-Baden, 37-48.

Billig, Michael. 2013. *Learn to Write Badly: How to Succeed in the Social Sciences*. Cambridge, UK: Cambridge University Press.

Bowers, Fredson. 1949. *Principles of Bibliographical Description*. Princeton, NJ: Princeton University Press.

Buckland, Michael. 2021. "Foreword". In *Hyperdocumentation* by Olivier Le Deuff, x-xiv. Hoboken, NJ: John Wiley & Sons.

Committee on Scientific and Technical Communication. 1969. *Scientific and Technical Communication. A Pressing National Problem and Recommendations for Its Solution*. Washington, DC: National Academy of Sciences.

Cowley, John Duncan. 1939. *Bibliographical Description and Cataloging*. London: Grafton & Co.

Cronin, Blaise. 2002. "Holding the Center while Prospecting at the Periphery: Domain Identity and Coherence in North American Information Studies Education". *Education for Information* 20, no. 1: 3–10. https://doi.org/10.3233/EFI-2002-20102

Egan, Margaret E. and Jesse H. Shera. 1952. "Foundations of a Theory of Bibliography". *Library Quarterly* 22, no. 2: 125-37.

Fjordback Søndergaard, Trine, Jack Andersen and Birger Hjørland. 2003. "Documents and the Communication of Scientific and Scholarly Information: Revising and Updating the UNISIST Model". *Journal of Documentation* 59, no. 3: 278–320.

Foot, Mirjam M. 2006. "The Study of Books". *Aslib Proceedings* 58, nos. 1-2: 20-33. DOI https://doi.org/10.1108/00012530610648653

de Fremery, Wayne de. 2024. *Cats, Carpenters, and Accounts. Bibliographical Foundations of Information Science.* The MIT Press, Cambridge, MA.

Galvin, Thomas J. 1977. "Pittsburgh. University of Pittsburgh Graduate School of Library and Information Sciences", In *Encyclopedia of Library and Information Science*, edited by Allen Kent, Harold Lancour and Jay E. Daily, 280-91. New York: Marcel Dekker.

Gants, David L. 2010. “Descriptive Bibliography and Electronic Publication”. In *Textual Cultures: Cultural Texts. Essays and Studies*, edited by Orietta Da Rold and Elaine M. Treharne. Cambridge, UK: Boydell & Brewer, 121-40. https://doi.org/10.1017/9781846158902.007

Gaskell, Philip. 1974. *A New Introduction to Bibliography*. Reprinted with corrections. Oxford: The Clarendon Press.

Greg, Walter Wilson. 1930. “The Present Position of Bibliography”. *The Library* 11, no. 3: 241-62.

Harter, Stephen P. 1986. *Online Information Retrieval. Concepts, Principles, and Techniques*. New York: Academic Press.

Henri, James and Lyn Hay. 1994. “Beyond the Bibliographic Paradigm: User Education in the Information Age”. *60th IFLA General Conference – Conference Proceedings – 21-27 August 1994,* La Habana, Cuba. https://origin-archive.ifla.org/IV/ifla60/60-henj.htm

Higgins, Julian P. T. and Sally Green, eds. 2009. The Cochrane Handbook for Systematic Reviews of Interventions, version 5.0.2. London: The Cochrane Collaboration. https://training.cochrane.org/handbook/archive/v5.0.2/

Hjørland, Birger and Lykke Kyllesbech Nielsen. 2001. "Subject Access Points in Electronic Retrieval". *Annual Review of Information Science and Technology* 35: 249- 98.

Hjørland, Birger. 1992. “The Concept of ‘Subject’ in Information Science”. *Journal of Documentation, 48, no. 2*: 172-200. https://doi.org/10.1108/eb026895

Hjørland, Birger. 2000. "Documents, Memory Institutions and Information Science". *Journal of Documentation* 56, no. 1: 27-41. https://doi.org/10.1108/EUM0000000007107

Hjørland, Birger. 2007. “Arguments for 'the Bibliographical Paradigm': Some Thoughts Inspired by the new English Edition of the UDC”. *Information Research. An International Electronic Journal* 12, no. 4. https://informationr.net/ir/12-4/colis/colis06.html

Hjørland, Birger. 2012. ”Om problemformuleringer i biblioteks-, dokumentations- og informationsvidenskab”. *Nordisk tidsskrift for informationsvidenskab og kulturformidling* 1, no. 3: 63-74. https://doi.org/10.7146/ntik.v1i3.26077.

Hjørland, Birger. 2013. “Information Science and its Core Concepts: Levels of Disagreement”. In *Theories of Information, Communication and Knowledge: A Multidisciplinary Approach*, edited by Fidelia Ibekwe-SanJuan and Thomas M. Dousa, 205-35. Dordrecht, the Netherlands: Springer. https://doi.org/10.1007/978-94-007-6973-1_9

Hjørland, Birger. 2015. “Classical Databases and Knowledge Organization: A Case for Boolean Retrieval and Human Decision-Making during Searches”. *Journal of the Association for Information Science and Technology* 66, no. 8: 1559-75. https://doi.org/10.1002/asi.23250

Hjørland, Birger. 2016. “Knowledge Organization”. *Knowledge Organization* 43, no. 6: 475-84. https://doi.org/10.5771/0943-7444-2016-6-475

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

Hjørland, Birger. 2018a. “Library and Information Science (LIS). Part 1”. *Knowledge Organization* 45, no. 3: 232-54. https://doi.org/10.5771/0943-7444-2018-3-232.

Hjørland, Birger. 2018b. “Library and Information Science (LIS). Part 2”. *Knowledge Organization* 45, no. 4: 319-38. https://doi.org/10.5771/0943-7444-2018-4-319.

Hjørland, Birger. 2023. “Bibliographical Control”. *Knowledge Organization* 50, no. 4: 301-15: https://doi.org/10.5771/0943-7444-2023-4-301

Hjørland, Birger. 2024a. “Bibliographical Foundations of Information Science: A Review Essay”. *Journal of Documentation*. Publication ahead-of-print, September 11, 2024. https://doi.org/10.1108/JD-06-2024-0126

Hjørland, Birger. 2024b. “Social Epistemology”. *Knowledge Organization* 51, no. 3: 187-202. https://doi.org/10.5771/0943-7444-2024-3-187

Kline, Ronald R. 2004. “What Is Information Theory a Theory Of? Boundary Work among Information Theorists and Information Scientists in the United States and Britain during the Cold War”. In *The History and Heritage of Scientific and Technical Information Systems: Proceedings of the 2002 Conference, Chemical Heritage Foundation*, edited by W. Boyd Rayward and Mary Ellen Bowden. Medford, NJ: Information Today, 15-28.

Krummel, D.W. 1986. *Bibliographies: Their Aims and Methods*. 2nd ed. London, UK: Mansell.

Krummel, D.W. 2017. “Bibliography”. In *Encyclopedia of Library and Information Sciences*, edited by John D. McDonald and Michael Levine-Clark, 468-79. 4th ed, vol. 1. Boca Raton, FL: CRC Press, vol. 1: 468-479.

McKenzie, D. F. 1999. *Bibliography and the Sociology of Texts*. Cambridge: Cambridge University Press.

Niro, Raymond, Beverley Crane and Louise Dagit, eds. 2008 *Dialog® Database Selection Guide*. 2008. [North Carolina]: The Thomson Corporation. http://web.archive.org/web/20240613193512/https://tefkos.comminfo.rutgers.edu/Courses/e530/Readings/Dialog_dbselguide.pdf

Otlet, Paul. 1990a. “Something about Bibliography”. In *International Organisation and Dissemination of Knowledge: Selected Essays of Paul Otlet*, translated and edited by W. Boyd Rayward, 11-24. Amsterdam: Elsevier.

Otlet, Paul. 1990b. “The Science of Bibliography and Documentation”. In *The International Organization and Dissemination of Knowledge: Selected essays of Paul Otlet*,

translated and edited by W. Boyd Rayward, 71-86. Amsterdam: Elsevier, 71–86.

President's Science Advisory Committee (1963). *Science, Government, and Information: the responsibilities of the technical community and the government in the transfer of information*. Washington, D.C.: The White House. http://garfield.library.upenn.edu/papers/weinbergreport1963.pdf

Rayward, W. Boyd. 2024. "The Attempt to Create a Post-World War II International Information Order". *IFLA Journal*. Publication ahead-of-print, May 14, 2024. https://doi.org/10.1177/03400352241242123

Reimer, Stephen R. 2015. "I.ii. Four Branches of Bibliographical Study". Manuscript Studies. Medieval and Early Modern (website). https://sites.ualberta.ca/~sreimer/ms-course/course/branchs.htm

Serrai, Alfredo. 1988-2001. *Storia della Bibliografia*. 11 vols. Rome: Bulzoni.

Shera, Jesse H. 1951. "Classification as the Basis of Bibliographic Organization." In *Bibliographic Organization: Papers presented before the Fifteenth Annual Conference of the Graduate Library School July 24-29, 1950*, edited by Jesse H. Shera and Margaret E. Egan, Chicago: Univ. of Chicago Press, 72-93.

Smiraglia, Richard P. 2001. "Further Progress toward Theory in Knowledge Organization". *The Canadian Journal of Information and Library Science* 26, nos. 2-3: 30–49. https://lazykoblog.knoworg.org/wp-content/uploads/2013/06/1_multipart_x005F_xf8ff_2_smiraglia.pdf

Smiraglia, Richard P. 2017. "Data and Metadata Instantiation: Use Cases and a Conceptual model". *In DCMI'17: Proceedings of the 2017 International Conference on Dublin Core and Metadata Applications 26-29 October 2017 Washington, DC, USA,* edited by Stuart A. Sutton and Paul Walk. Washington, DC: Dublin Core Metadata Initiative, 25-26. https://dl.acm.org/doi/abs/10.5555/3183831.3183836

Smiraglia, Richard P. and Charles van den Heuvel. 2013. "Classifications and Concepts: Towards an Elementary Theory of Knowledge Interaction". *Journal of Documentation* 69, no. 3: 360-83. https://doi.org/10.1108/JD-07-2012-0092

Tanselle, G. Thomas. 1977. "Descriptive Bibliography and Library Cataloguing". *Studies in Bibliography* 30: 1-56.

Tanselle, G. Thomas. 1990. *Textual Criticism and Scholarly Editing*. Charlottesville: Bibliographical Society of the University of Virginia.

The Times Comprehensive Atlas of the World. 2014. 14th ed. London: Times Books.

UNESCO (United Nations Educational, Scientific and Cultural Organization) and ICSU (International Council of Scientific Unions). 1971. *UNISIST: Study Report on the Feasibility of a World Science Information System*. Paris: UNESCO.

Van Acker, Wouter. 2011. "Universalism as Utopia : A Historical Study of the Schemes and Schemas of Paul Otlet (1868-1944)". PhD diss., Ghent University.

Webb, William H., Alan R. Beals and Carl Milton White, eds. 1986. *Sources of Information in the Social Sciences. A Guide to the Literature*. 3rd ed. Chicago: American Library Association.

Yee, Martha M. 2007. "Cataloging Compared to Descriptive Bibliography, Abstracting and Indexing Services, and Metadata". *Cataloging & Classification Quarterly* 44, nos. 3–4: 307–27. https://doi.org/10.1300/J104v44n03_10

Peer Review in 2024

Knowledge Organization, like all scientific journals, relies on extensive peer review of manuscript submissions to ensure the originality and quality of research published in our pages. On behalf of the KO Editorial Board we would like to express our unqualified gratitude to the scholars listed below who provided peer review during 2024.

Aida Slavic
Benildes Maculan
Beth Patin
Bidyarthi Dutta
Birger Hjørland
Bruno Almeida
Carlos Henrique Marcondes
Claudia Lanza
Claudio Gnoli
Daniel Martínez-Ávila
David Haynes
Fidelia Ibekwe
Jane Greenberg
Jonathan Furner
José Augusto Chaves Guimarães
Julia Bullard
Linair Maria Campos
Lu An
M.P. Satija
Maja Žumer,
Marcia Zheng
Mayukh Bagchi
Michely Jabala Mamede Vogel
Muhammad Rosyihan Hendrawan
Pat Riva
Preedip B Balaji
Rexwhite Tega Enakrire
Richard P. Smiraglia
Rick Szostak
Suellen Milani
Sylvie Davis
Thiago Barros
Thomas Dousa
Vanda Broughton
Widad Mustafa El Hadi

KNOWLEDGE ORGANIZATION KO

Official Journal of the International Society for Knowledge Organization ISSN 0943 – 7444

International Journal devoted to Concept Theory, Classification, Indexing and Knowledge Representation

Publisher

Ergon – ein Verlag in der Nomos Verlagsgesellschaft mbH
Waldseestraße 3-5
D-76530 Baden-Baden
Tel. +49 (0)7221-21 04-667
Fax +49 (0)7221-21 04-27
Sparkasse Baden-Baden Gaggenau
IBAN: DE05 6625 0030 0005 0022 66
BIC: SOLADES1BAD

KNOWLEDGE ORGANIZATION
Journal of the International Society for Knowledge Organization
Natália B. Tognoli, Editor-in-Chief
nataliatognoli@id.uff.br

Instructions for Authors

Manuscripts should be submitted electronically (in Microsoft® Word format) in English only via Open Journal System at https://ojs.nomos-journals.de/index.php/ko . Manuscripts that do not adhere to these guidelines will be returned to the authors in proper form for resubmission.
Manuscripts should be accompanied by an indicative abstract of approximately 250 words. Manuscripts of articles should fall within the range of 6,000-10,000 words. Longer manuscripts will be considered on consultation with the Editor-in-chief.

A separate title page should include the article title and the author's name, postal address, and E-mail address. Only the title of the article should appear on the first page of the text. Contact information must be present for all authors of a manuscript.

Please consider suggesting potential reviewers for the manuscript and provide their e-mail addresses. Potential reviewers should have published in the area of research and must not be from the same institutions as the authors or have collaborated with them in the past.

The author's name should not appear on the manuscript to protect anonymity.

Criteria for acceptance will be appropriateness to the field of knowledge organization (see "Scope" at https://www.isko.org/ko.html), taking into account the merit of the contents and presentation. It is expected that all successful manuscripts will be well-situated in the domain of knowledge organization and will cite all relevant literature from within the domain. Authors are encouraged to use the KO literature database at http://www.isko.org/lit.html.

The manuscript should be concise and should conform to professional standards of English useage and grammar. Authors whose native language is not English are encouraged to make use of professional academic English-language proofreading services. We recommend Vulpine Academic Services (vulpineacademic@gmail.com).

Manuscripts are received with the understanding that they have not been previously published, are not being submitted for publication elsewhere, and that if the work received official sponsorship, it has been duly released for publication. Submissions are refereed, and authors will usually be notified within 6 to 8 weeks.

Manuscript Formatting

Under no circumstances should the author attempt to mimic the presentation of text as it appears in our published journal. Instead, please follow these instructions:

In Microsoft® Word please set the language preference ("Tools," "Language") to "English (US)" or "English (UK)."

The entire manuscript should be double-spaced, including notes and references.

The text should be structured with decimally-numbered subheadings (1.0, 1.1, 2.0, 2.1, 2.1.1, etc.). It should contain an introduction, giving an overview and stating the purpose, a main body, describing in sufficient detail the materials or methods used and the results or systems developed, and a conclusion or summary.

Author-generated keywords are not permitted.

Footnotes are not allowed but endnotes may be used. Do not use automatic footnote formatting. Instead, insert a superscript numeral (Format, Font, Superscript) and create the text of the note manually in a separate list at the end of the manuscript, before the reference list.

Paragraphs should include a topic sentence, a developed narrative and a conclusion; a typical paragraph has several sentences. Paragraphs with tweet-like characteristics (one or two sentences) are inappropriate.

Italics are permitted only for phrases from languages other than English and for the titles of published works.

Bold type is not permitted.

Em-dashes should not be used as substitutes for commas. Dashes must be inserted manually (Insert, Advanced Symbol, Em-dash) with no spaces on either side.

Illustrations should be embedded within the document. Photographs (including color and halftone) should be scanned with a minimum resolution of 600 dpi and saved as .jpg files. Tables should contain a number and caption at the bottom, and all columns and rows should have headings. All illustrations should be cited in the text as Figure 1, Figure 2, etc. or Table 1, Table 2, etc.

Examples of KOS (classification, taxonomy, ontology, thesaurus, etc.) arrays should be configured as figures and set into the document as jpg; they should not be entered as editable text.

Remove all active hyperlinks, including those from reference formatting software (if hovering over the text with a mouse produces a gray highlight, the text is hyperlinked; remove the link "Insert," "Hyperlink," "Remove link").

Upon acceptance of a manuscript for publication, authors must provide a digital photo and a one-paragraph biographical sketch (fewer than 100 words). The photograph should be scanned with a minimum resolution of 600 dpi and saved as a .jpg file.

References

Reference citations within the text should have the form: (Author year). For example, (Jones 1990). Specific page numbers are required for quoted material, e.g. (Jones 1990, 100). A citation with two authors would read (Jones and Smith 1990); three or more authors would be: (Jones et al. 1990). When the author is mentioned in the text, only the date and optional page number should appear in parentheses: "According to Jones (1990), ..." or "Smith wrote (2010, 146): ...." A subsequent page reference to the same cited work (e.g., to Smith 2010) should have the form "(229)." There is never a comma before the date.

In-text citations should not be routinely placed at the end of a sentence or after a quotation, but an attempt should be made to work them into the narrative. For example:

Jones (2010, 114) reported statistically significant results.

Many authors report similar data; according to Matthews (2014, 94): "all seven studies report means within ±5%."

In-text citations should precede block quotations and are never placed at the end of a block quotation.

References should be listed alphabetically by author at the end of the article. Reference lists should not contain references to works not cited in the text. Websites mentioned in passing in the text should be identified parenthetically with their URLs but not with references unless a specific page of a specific website is being quoted.

Author names should be given as found in the sources (not abbreviated, but also not fuller than what is given in the source). Journal titles should not be abbreviated. Multiple citations to works by the same author should be listed chronologically and should each include the author's name. Articles appearing in the same year should have the following format: "Jones 2005a, Jones 2005b, etc."

Proceedings must be identified fully by volume title (which often differs from the name of the conference), editor(s), series (if applicable), and details of publication (place and publisher name). Pa- pers in proceedings must be identified by their page numbers. Papers from online proceedings that lack publication details should include a DOI or URL.

References for published electronic resources should be accompanied by either a URL or DOI but not in lieu of actual publication data; access dates are not allowed.

Unpublished electronic resources may use an access date in lieu of the date of publication. In cases of doubt, authors are encouraged to consult *The Chicago Manual of Style* 17th ed. (or online), author-date reference system (chapter 15).

Examples:

Books

Ingwersen, Peter. 1992. *Information Retrieval Interaction*. London: Taylor Graham.

Mayr, Ernst. 1991. *Principles of Systematic Zoology*, 2nd ed. New York: McGraw-Hill.

Chapters

Egan, Frances. 2012. "Representationalism". In *Oxford Handbook of Cognitive Science*, edited by Eric Margolis, Richard Samuels, and Stephen P. Stich, 250-72. Oxford: Oxford University Press.

Articles

Furner, Jonathan. 2007. "Dewey Deracialized: A Critical Race-Theoretic Perspective". *Knowledge Organization* 34, no. 3: 144–68. https://doi.org/10.5771/0943-7444-2007-3-144

Schallier, Wouter. 2004. "On the Razor's Edge: Between Local and Overall Needs in Knowledge Organization." *Advances in Knowledge Organization* 9: 269-74.

Proceedings:

Sameh S. El-Atawy, and Mohamed E. Khalefa. 2016. "Building an Ontology-Based Electronic Health Record System". In *Proceedings of the 2nd Africa and Middle East Conference on Software Engineering (AMECSE '16) 28 – 29 May 2016 Cairo, Egypt*. Association for Computing Machinery, New York, NY, USA, 40–45. https://doi.org/10.1145/2944165.2944172

PhD dissertation

Olson, Hope Alene. 1996. "The Power to Name: Marginalizations and Exclusions of Subject Representation in Library Catalogues". PhD diss., University of Wisconsin

KO is published by Ergon.

Annual subscription 2024:
- **Print + online (8 issues/ann.; unlimited access for your Campus via Nomos eLibrary) € 429,00/ann.**
- **Prices do not include postage and packing**
- **Cancellation policy: Termination within 1 month notice to the end of the calendar year**

KO KNOWLEDGE ORGANIZATION

Official Journal of the International Society for Knowledge Organization ISSN 0943 – 7444

International Journal devoted to Concept Theory, Classification, Indexing and Knowledge Representation

Scope

The more scientific data is generated in the impetuous present times, the more ordering energy needs to be expended to control these data in a retrievable fashion. With the abundance of knowledge now available the questions of new solutions to the ordering problem and thus of improved classification systems, methods and procedures have acquired unforeseen significance. For many years now they have been the focus of interest of information scientists the world over.

Until recently, the special literature relevant to classification was published in piecemeal fashion, scattered over the numerous technical journals serving the experts of the various fields such as:

philosophy and science of science
science policy and science organization
mathematics, statistics and computer science
library and information science
archivistics and museology
journalism and communication science
industrial products and commodity science
terminology, lexicography and linguistics

Beginning in 1974, KNOWLEDGE ORGANIZATION (formerly INTERNATIONAL CLASSIFICATION) has been serving as a common platform for the discussion of both theoretical background questions and practical application problems in many areas of concern. In each issue experts from many countries comment on questions of an adequate structuring and construction of ordering systems and on the problems of their use in opening the information contents of new literature, of data collections and survey, of tabular works and of other objects of scientific interest. Their contributions have been concerned with

(1) clarifying the theoretical foundations (general ordering theory/science, theoretical bases of classification, data analysis and reduction)
(2) describing practical operations connected with indexing/classification, as well as applications of classification systems and thesauri, manual and machine indexing
(3) tracing the history of classification knowledge and methodology
(4) discussing questions of education and training in classification
(5) concerning themselves with the problems of terminology in general and with respect to special fields.

Aims

Thus, KNOWLEDGE ORGANIZATION is a forum for all those interested in the organization of knowledge on a universal or a domain-specific scale, using concept-analytical or concept-synthetical approaches, as well as quantitative and qualitative methodologies. KNOWLEDGE ORGANIZATION also addresses the intellectual and automatic compilation and use of classification systems and thesauri in all fields of knowledge, with special attention being given to the problems of terminology.

KNOWLEDGE ORGANIZATION publishes original articles, reports on conferences and similar communications, as well as book reviews, letters to the editor, and an extensive annotated bibliography of recent classification and indexing literature.

KNOWLEDGE ORGANIZATION should therefore be available at every university and research library of every country, at every information center, at colleges and schools of library and information science, in the hands of everybody interested in the fields mentioned above and thus also at every office for updating information on any topic related to the problems of order in our information-flooded times.

KNOWLEDGE ORGANIZATION was founded in 1973 by an international group of scholars with a consulting board of editors representing the world's regions, the special classification fields, and the subject areas involved. From 1974-1980 it was published by K.G. Saur Verlag, München. Back issues of 1978-1992 are available from ERGON-Verlag, too.

As of 1989, KNOWLEDGE ORGANIZATION has become the official organ of the INTERNATIONAL SOCIETY FOR KNOWLEDGE ORGANIZATION (ISKO) and is included for every ISKO-member, personal or institutional in the membership fee.

Annual subscription 2024: Print + online (8 issues/ann.; unlimited access for your Campus via Nomos eLibrary) € 429,00/ann. Prices do not include postage and packing. Cancellation policy: Termination within 1 month notice to the end of the calendar year

Ergon – ein Verlag in der Nomos Verlagsgesellschaft mbH, Waldseestraße 3-5, D-76530 Baden-Baden, Tel. +49 (0)7221-21 04-667, Fax +49 (0)7221-21 04-27, Sparkasse Baden-Baden Gaggenau, IBAN: DE05 6625 0030 0005 0022 66, BIC: SOLADES1BAD

Founded under the title International Classification in 1974 by Dr. Ingetraut Dahlberg, the founding president of ISKO. Dr. Dahlberg served as the journal's editor from 1974 to 1997, and as its publisher (Indeks Verlag of Frankfurt) from 1981 to 1997.

The contents of the journal are indexed and abstracted in *Social Sciences Citation Index, Web of Science, Information Science Abstracts, INSPEC, Library and Information Science Abstracts* (LISA), *Library, Information Science & Technology Abstracts* (EBSCO), *Library Literature and Information Science* (Wilson), *PASCAL, Referativnyi Zhurnal Informatika, and Sociological Abstracts.*