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

Terminological Visibility of LGBTQIAPN+ Themes in Knowledge Organization

Raphael Augusto dos Santos
Federal University of São Carlos (UFSCar), Brazil

Maria Laura Machado Cangiani
Federal University of São Carlos (UFSCar), Brazil

Lucas Gatto Covello
São Paulo State University "Júlio de Mesquita Filho" (UNESP), Brazil

Paula Regina Dal'Evedove
Federal University of São Carlos (UFSCar), Brazil

Leandro Innocentini Lopes de Faria
Federal University of São Carlos (UFSCar), Brazil

Abstract

This study analyzes how LGBTQIAPN+ themes are represented in the keywords of scientific output in Knowledge Organization indexed in the Scopus and Web of Science databases. The investigation understands keywords as non-neutral descriptors, as metadata that participate in the naming, retrieval, and visibility of historically marginalized subjects and experiences. Methodologically, it adopts a mixed-methods approach, combining bibliometric procedures and qualitative interpretation of terminological patterns. Recent growth in the theme has been observed, but with unequal visibility across the databases. In Scopus, there is a greater presence of identity terms, such as queer, gender, transgender, and LGBT, articulated with thematic representation practices. In Web of Science, the disciplinary lexicon of Knowledge Organization stands out, such as classification, subject access, and representation. The study concludes that the terminological visibility of LGBTQIAPN+ remains partial, relational, and dependent on the forms of association established in scientific metadata.

1.0 Introduction

Knowledge Organization Systems are products of specific historical contexts, and their hierarchical structures can subordinate, disperse, erase, or marginalize themes, producing access problems (Martínez-Ávila and Fox 2025). By selecting concepts, authorizing terms, establishing categories, and defining semantic hierarchies, Knowledge Organization participates in the construction of the conditions by which subjects, experiences, and knowledge domains become visible, retrievable, and legitimized in information systems.

This critical dimension becomes relevant in the face of historically marginalized memories, subjectivities, and identities. Representation languages can both broaden access and recognition and reproduce asymmetries, prejudices, and exclusions in the instruments and

metadata that organize information. Critical literature in the field has demonstrated that naming is also an exercise of power, in which Knowledge Organization Systems delimit what can appear, under which categories, and in what position it will be situated within an order of knowledge. Olson (2002), in discussing the power of naming, shows that controlled vocabularies and representation systems, while claiming universality, produce limits through the inclusions and exclusions they perform. Thematic representation not only describes the documentary universe but also participates in the construction of conceptual frameworks through which subjects, identities, and experiences become searchable, recognizable, and intelligible.

This issue takes on particular complexity in the LGBTQIAPN+ domain, traversed by terminological historicity, political disputes over naming, semantic mobility, and internal heterogeneity. The terms used to represent LGBTQIAPN+ subjects and experiences may carry marks of pathologization, stigmatization, and erasure, but they can also express processes of recognition, self-definition, and identity affirmation. The classification of queer phenomena is marked by a constitutive paradox, in which categories are necessary to enable institutional access, yet at the same time they can regulate, stabilize, and reinscribe subjects within recognizable boundaries (Keilty 2009). The terminological insufficiency in informational infrastructures compromises the findability of queer and trans materials, while the rapid transformation of LGBTQIAPN+ language challenges static vocabularies and rigid hierarchical structures (Cifor and Rawson 2023).

In this research, terminological visibility is understood as the capacity of certain subjects, identities, experiences, and debates to become nameable, retrievable, and semantically articulable through the terms employed in the metadata of scientific communication. This notion allows us to shift the analysis from the simple presence or absence of terms to an understanding of the ways in which they are selected, combined, and associated in specific contexts of representation. Although the critical literature on Knowledge Organization discusses the representation of LGBTQIAPN+ subjects and experiences, further investigation is needed into how the scientific production of the field itself names and articulates these themes in its metadata, especially in keywords.

From this perspective, this research investigates how LGBTQIAPN+ themes are represented in the keywords within scientific production in Knowledge Organization indexed in the Web of Science and Scopus databases. More specifically, it seeks to identify naming patterns, thematic association, and terminological visibility. The contribution of this research lies in shifting the analysis of LGBTQIAPN+ representation from Knowledge Organization Systems to the metadata of scientific communication in the field, highlighting how keywords contribute to the visibility, recognition, and sense of belonging of the LGBTQIAPN+ community within Knowledge Organization.

2.0 Subject Representation, Terminology and LGBTQIAPN+ Themes in Knowledge Organization

In the LGBTQIAPN+ domain, terminology cannot be understood merely as a stable set of descriptors, since labels and terminologies incorporate ideological assumptions and affect who obtains recognition, support, and community protection (Thelwall et al. 2023). In the academic context, these terms are also mobilized to delimit themes and objects addressed in scientific research, requiring researchers to choose keywords that are simultaneously precise, appropriate, and sensitive to the transformations of social language. Terminological standardization, while favoring information retrieval, can also reject identities, become obsolete in the face of linguistic changes, and produce groups excluded from the adopted definitions (Eliason 2014).

This terminological instability highlights the limitations of Knowledge Organization Systems when confronted with domains marked by more intense historical and sociocultural transformations. Fujita et al. (2012) draw attention to the obsolescence and outdated nature of thematic representation instruments, arguing that these should keep pace with the evolution of terminology and conceptual relationships established in each domain and sociocultural context. This discussion is particularly relevant when considering issues such as gender and sexuality. Santana (2014) observes that these themes have gained academic momentum recently: gender, through the strengthening of feminist debates in the second half of the 20th century; and sexuality, through the intensification of studies from the 1980s onwards. This means that many Knowledge Organization Systems were structured before the consolidation of these debates and their conceptual, political, and social developments.

Butler (2014) questions restrictive understandings of gender based exclusively on biological foundations and proposes that social, psychological, hormonal, interstitial, and performative dimensions should also be considered. This shift highlights that domains related to gender and sexuality cannot be reduced to fixed and binary classifications. It is in this context that queer studies become relevant for reflecting on the representation of information. Santana (2014) observes that the term queer refers to what is strange, different, eccentric, ambiguous, changeable, fluid, and multiple. Pinho (2010), by listing subgroups such as lesbians, transvestites, transsexuals, transgender people, bisexuals, drag queens, cross-dressers, intersex people, asexuals, and allies, reinforces the internal plurality of the domain and the need to represent it with greater specificity. A direct consequence of this complexity is the challenge of constructing forms of representation that are sufficiently broad and, at the same time, conceptually adequate. In this regard, adequate representation is not only a technical requirement of the system, but also an ethical condition for access to information.

This complexity can also be understood from the tensions between categorization, identity, and intersectionality in Knowledge Organization. Martínez-Ávila and Fox (2025) argue that identities such as race, gender, class, national origin, and disability cannot be treated as isolated categories, as they mutually constitute each other and produce intersectional forms of experience and oppression. For the authors, this debate is central, as classificatory systems operate on categories of human groups and can reproduce problems such as essentialism, erasure, dispersion, and subordination. Representing LGBTQIAPN+ subjects and themes requires recognizing that identity categories are

necessary for visibility and information retrieval, but can also fix, reduce, or fragment complex social experiences. The effects of terminological inadequacy become visible when considering the ways in which LGBTQIAPN+ subjects have been historically named. Moura (2018) demonstrates that the terms attributed to homosexual subjects have, at different times, been associated with pathologizing perspectives, as can be seen in the use of designations such as "inverted" and in the historical association between homosexuality and illness.

This terminological dispute also involves the question of who has the authority to name subjects, practices, and identities. Lo Vecchio (2026), in problematizing the LGBTQ lexicon in the field of historical linguistics, argues that the researcher's position in relation to the object of study cannot be dissociated from the words used to name it, especially when dealing with socially marked themes such as gender and sexuality. For the author, authority over language is also a matter of power, in which lexical choices both reinforce and contest ideologies and social relations. By drawing on McConnell-Ginet, Lo Vecchio (2026) affirms that communities constitute the ultimate semantic authorities, which reinforces the need to consider the ways in which LGBTQIAPN+ subjects and collectivities produce, negotiate, and claim their own terms of identification.

Marginalized individuals and communities produce their own repertoires of identification. Self-naming constitutes a form of empowerment, through which subordinate communities claim the possibility of naming themselves, their experiences, and their modes of belonging (Campbell et al. 2017). From this perspective, terms not only describe previously given identities but also participate in the construction of collective meanings, functioning as instruments of affirmation, resistance, protection, and recognition. For Knowledge Organization, this debate is central as it highlights that representing LGBTQIAPN+ individuals requires considering not only institutional terminological correctness but also the vocabularies produced, negotiated, and legitimized by the communities themselves.

The identity dimension of naming can also be observed in the experiences of transgender people with the choice of their own names. Pamfile et al. (2025), in investigating renaming processes among transgender people, indicate that the use of a name congruent with gender identity can favor self-esteem and well-being, in addition to performing classificatory, identificatory, and symbolic functions. The authors also show that the chosen name can express both continuity and rupture with the past, being experienced as a metaphor for a new beginning, support for continuity through transformation, or compromise between family and cultural identifications and disidentifications. Although situated in the field of mental health and trans experiences, this study contributes to understanding that naming and self-naming are not merely linguistic acts, but practices of identity affirmation, social recognition, and reconstruction of belonging.

Henry et al. (2022), by examining the representation of sex, sexuality, and queerness in the Library of Congress Classification, argue that the problems are not limited to the authorized terms in controlled vocabularies, but extend to the classification structure itself. For the authors, the classification functions as a kind of "skeleton" of bibliographic organization, upon which controlled vocabularies are based; therefore, simply changing the headings

does not produce effective change when the structure continues to reproduce erasures, stigmatizations, harmful associations, and forms of disempowerment. The disputes surrounding naming show that classification does not only act upon a previously given social reality; it also participates in the production of the ways in which certain groups can be understood, located, and legitimized.

As a practical response to these limitations, specialized controlled vocabularies have been adopted to broaden LGBTQIAPN+ materials representation and retrieval. Fischer (2023) notes that the Library of Congress Subject Headings on LGBTQIA+ topics have not evolved at the same pace as the vernacular vocabulary used by communities and users when searching for informational resources. In this context, Homosaurus: An International LGBTQ+ Linked Data Vocabulary is presented as an alternative controlled vocabulary, capable of complementing the LCSH when its terms do not adequately represent the cataloged materials. Being a specialized, free, online-available, and periodically updated vocabulary, Homosaurus constitutes a concrete response to the limitations of general vocabularies, favoring greater alignment between document representation, community language, and contemporary search practices.

In scientific communication, keywords play a relevant role as thematic representation metadata, functioning as access points that guide the retrieval, indexing, circulation, and aggregation of studies in databases. Although they can be assigned by the authors themselves, these words do not operate outside of disciplinary conventions, editorial policies, and academic visibility strategies. Thus, their selection participates in the construction of the ways in which certain research objects become retrievable, recognizable, and semantically associated within a scientific field. When it comes to LGBTQIAPN+ themes, this issue becomes even more sensitive. The choice of keywords can broaden the visibility of historically marginalized groups, identities, and experiences, but it can also produce information barriers when it resorts to generic, outdated terms or terms poorly aligned with contemporary forms of naming. From this perspective, analyzing the keywords of scientific production in Knowledge Organization allows for the observation not only which terms appear, but also how the field itself selects, stabilizes, combines, and makes LGBTQIAPN+ themes retrievable through the metadata that organizes its scientific communication.

3.0 Methodology

To achieve this objective, a mixed-methods approach was adopted, combining bibliometric procedures and qualitative interpretation of the terminological patterns present in the metadata of scientific production. The quantitative stage was employed to map the temporal distribution, frequency, and co-occurrence of keywords related to the LGBTQIAPN+ domain in Knowledge Organization studies. The qualitative stage consisted of a critical interpretation of the identified naming patterns and thematic association. Data collection was carried out on February 25, 2026, in the Scopus and Web of Science databases, selected for their international coverage and relevance in bibliometric studies. The search was carried out without time frame, using the expressions (“LGBT” OR “LGBTQ”

OR “LGBTI” OR “queer” OR “transgender” OR “transsexual” OR “nonbinary” OR “non-binary” OR “sexual minority” OR “gender minority” OR “gender identity” OR “sexual orientation”) AND (“knowledge organization” OR “knowledge organisation”). The adoption of the search strategy in English stems from a methodological need associated with the functioning of the consulted sources, and does not imply disregarding the relevance of scientific production published in other languages.

The inclusion of the forms “knowledge organization” and “knowledge organisation” encompasses orthographic variations of American and British English. It is recognized that this option limits the retrieval of studies that use local, regional, or non-English terminologies. The search retrieved 111 records in Scopus and 26 records in Web of Science. The databases were analyzed separately to compare the modes of terminological representation in each indexing environment. Records that articulated LGBTQIAPN+ themes with Knowledge Organization, thematic representation, classification, indexing, metadata, controlled vocabularies, folksonomies, ontologies, or related themes were considered. The exported metadata included title, abstract, keywords, year of publication, authorship, journal, and institutional data. The analysis focused on keywords, understood as thematic access points and metadata that participate in the construction of terminological visibility in scientific communication. The set of keywords was reviewed for graphic variations, singular and plural forms, acronyms, and semantically equivalent expressions.

The analyses were performed using Bibliometrix/Biblioshiny in the RStudio environment to organize the corpus and generate descriptive indicators, such as the annual distribution of publications; and VOSviewer to construct keyword co-occurrence networks and identify thematic clusters. In Web of Science, a minimum threshold of two occurrences per term was adopted; in Scopus, five occurrences. This difference stems from variations in the size and density of the corpora, aiming to produce interpretable and comparable networks. The networks were analyzed considering the frequency of the terms, their position, their connections, and the thematic composition of the clusters.

4.0 Results

Figure 1 shows the annual distribution of documents retrieved from the Scopus and Web of Science databases. The represented years correspond to the intervals actually retrieved in each database. The difference between the retrieved sets (111 records in Scopus and 26 in Web of Science) highlights the distinct coverage of the databases and helps to explain the higher terminological density observed in Scopus. It is observed that Scopus presents a more extensive time series, with documents distributed between 1989 and 2025, while Web of Science presents records between 1996 and 2026. The higher incidence of documents in Scopus may be related to its broader coverage in areas close to the Humanities and Social Sciences, as observed by Mongeon and Paul-Hus (2016). Production fluctuates throughout the period, but grows more significantly from 2016 onwards, reaching its highest numbers between 2024 and 2025. In the Web of Science, a more dispersed distribution is observed, with gaps between 2001 and 2012 and more

noticeable growth from 2022 onwards. The drop observed in 2026 should be interpreted with caution, since the data collection was carried out on February 25, 2026. In analytical terms, recent growth identified in both databases indicates a growing interest in themes related to the LGBTQIAPN+ community within the scope of scientific communication in Knowledge Organization, even though this presence is uneven and strongly conditioned by the coverage, indexing, and representation policies of each database.

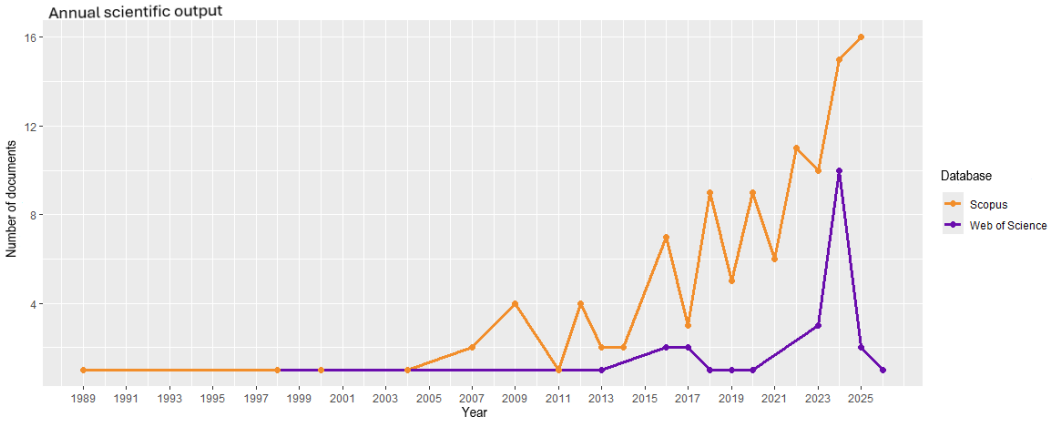


Figure 1. Annual distribution of documents retrieved from the Web of Science and Scopus databases, generated using Bibliometrix in RStudio.

Figure 2 presents the word clouds constructed from the most recurrent terms in the sets retrieved from each database.



Figure 2. Comparison of word clouds from the Scopus and Web of Science databases, generated using Biblioshiny.

In Scopus, there is a predominance of words directly associated with the thematic domain being investigated, such as queer, gender, and transgender, as well as terms related to information representation and organization, such as subject index term, knowledge organization and classification. This set suggests a more robust presence of identity vocabularies and sexual and gender diversity, indicating a thematic representation attentive to contemporary forms of naming this domain. In the Web of Science, in turn,

terms linked to the disciplinary lexicon of Knowledge Organization, such as knowledge organization, classification, subject access and representation become more prominent. Despite the presence of expressions related to gender and sexuality, these are less central in the word cloud when compared to other terms, suggesting that, in this environment, the LGBTQIAPN+ theme tends to be more strongly mediated by the conceptual categories of the field itself. This contrast between the databases indicates that terminological visibility is not distributed homogeneously. While Scopus highlights identity vocabularies more intensely, Web of Science more strongly emphasizes the disciplinary vocabulary of the area. In analytical terms, this suggests different forms of articulation between the LGBTQIAPN+ domain and the Organization of Knowledge, oscillating between a visibility more centered on subjects and identities and another more centered on the categories of the field.

To investigate deeper, keyword co-occurrence networks, presented in Figures 3 and 4, were analyzed. While word clouds indicate the relative frequency of terms, the networks allow for the observation of associations established between them, highlighting not only the presence of the terms but also the ways in which themes are articulated within the metadata.

Within the Scopus network, three clusters have emerged, in which terms such as queer, transgender, gender, and LGBTQ appear articulated both among themselves and with expressions associated with Knowledge Organization, such as knowledge organization, classification and subject indexing. This configuration suggests that, within this set, LGBTQIAPN+ themes do not appear merely as isolated topics, but as elements semantically connected to Knowledge Organization processes and instruments.

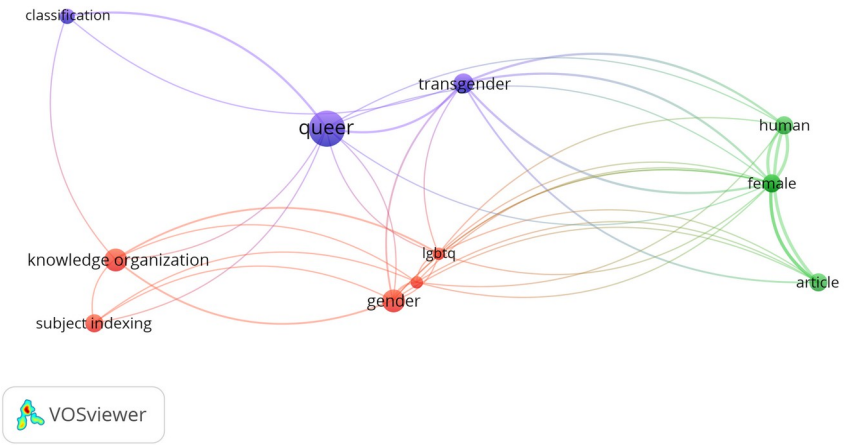


Figure 3. Keyword co-occurrence network on the theme in the Scopus database, generated using VOSviewer.

In the Web of Science, the network presents a more segmented structure. The term knowledge organization occupies a central position and is linked to expressions such as classification, representation, subject access and information science, highlighting the greater centrality of the disciplinary lexicon of the area. Expressions related to gender and sexuality, such as lgbtq, gender, women, and queer theory also appear, but distributed in clusters less integrated into the main core. In addition, a specific cluster associated with ontology and Wikidata is observed, indicating a more technical and infrastructural articulation.

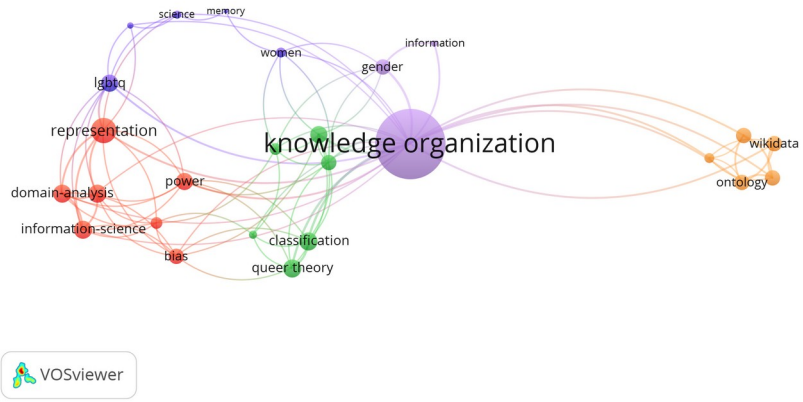


Figure 4. Keyword co-occurrence network on the theme in the Web of Science database, generated using VOSviewer.

In analytical terms, asymmetries between Scopus and the Web of Science reveal that the visibility of LGBTQIAPN+ themes reflects how concepts are stabilized in science. Scopus demonstrates greater openness to the contemporary language of diversity, while the Web of Science conditions the domain to the framework of the consolidated disciplinary lexicon of Knowledge Organization. This division proves that representation in global systems is relational and conditioned by regimes of terminological authority. This finding highlights that metadata is not neutral, but rather constructs the intelligibility of studies. In domains marked by linguistic mobility, the work of information professionals requires critical and ethical mediation focused on self-identification. Rigid systems insensitive to social transformations generate information barriers that limit the visibility of marginalized groups. Therefore, the disparity between the databases is not a mere variation in retrieval but reflects the degree of cultural warrant in the scientific communication of Knowledge Organization itself. Notably, the scientific community of Knowledge Organization must assign keywords in an equitable, inclusive, and socially responsible manner in research on LGBTQIAPN+.

5.0 Conclusion

This study highlights that the representation of LGBTQIAPN+ themes through keywords in scientific communications in Knowledge Organization is unequal and relational, constituting an expression of the disputes over terminological authority that permeate the construction of scientific metadata. Far from being a purely descriptive problem, this finding demonstrates that keywords operate as political devices which affect the intelligibility of historically marginalized experiences. The main contribution of the study lies in showing that the scientific community of Knowledge Organization itself reproduces, in its scientific metadata, the tensions between terminological institutionalization and the emergence of new forms of social naming. This tension becomes evident when the analyzed scientific metadata privilege a consolidated disciplinary lexicon over vernacular and identity-based terminology, potentially limiting the visibility of emerging forms of social naming. The findings reveal that the scientific community of Knowledge Organization should adopt a critical and sensitive stance when assigning metadata related to the LGBTQIAPN+ community, considering that keyword selection and assignment constitute a form of interference in naming practices in science, in which the choice of terms can act as an instrument of conceptual inclusion or exclusion. As limitations, the study was restricted to the Scopus and Web of Science databases, the English language, and keywords. For future research, the study suggests expanding the corpus to other databases and comparing the keywords assigned by scientific authors with controlled descriptors in order to map the tensions between authorial and institutional language.

AI Usage Statement

Generative artificial intelligence tools were used to support language revision, translation, and stylistic refinement of the manuscript. The conceptual development, methodological design, analysis, interpretation of results, verification of references, and final content were carried out and reviewed by the authors, who take full responsibility for the manuscript.

References

Butler, Judith. 2014. “Regulações de gênero.” Cadernos Pagu, no. 42: 249–274. https://doi.org/10.1590/0104-8333201400420249.

Campbell, D. Grant, José Augusto Chaves Guimarães, Fabio Assis Pinho, Daniel Martínez-Ávila, and Francisco Arrais Nascimento. 2017. “The Terminological Polyhedron in LGBTQ Terminology: Self-Naming as a Power to Empower in Knowledge Organization.” Knowledge Organization 44, no. 8: 586–591. https://doi.org/10.5771/0943-7444-2017-8-586.

Cifor, Marika, and K. J. Rawson. 2023. “Mediating Queer and Trans Pasts: The Homosaurus as Queer Information Activism.” Information, Communication & Society 26, no. 11: 2168–2185. https://doi.org/10.1080/1369118X.2022.2072753.

Eliason, Michele J. 2014. “An Exploration of Terminology Related to Sexuality and Gender: Arguments for Standardizing the Language.” Social Work in Public Health 29, no. 2: 162–175. https://doi.org/10.1080/19371918.2013.775887.

Fischer, Rachel K. 2023. “Using the Homosaurus in a Public Library Consortium: A Case Study.” Library Resources & Technical Services 67, no. 1: 4–15. https://journals.ala.org/index.php/lrts/article/view/7985/11110.

Fujita, Mariângela Spotti Lopes, María del Carmen Agustín Lacruz, and Raquel Gómez Díaz. 2012. “A situação atual da indexação nas tarefas bibliotecárias.” Perspectivas em Ciência da Informação 17, no. 1: 94–109. https://doi.org/10.1590/S1413-99362012000100006.

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, June 1. https://www.inthelibrarywiththeleadpipe.org/2022/the-old-and-the-prudish/.

Keilty, Patrick. 2009. “Tabulating Queer: Space, Perversion, and Belonging.” Knowledge Organization 36, no. 4: 240–249. https://www.imrpress.com/journal/KO/36/4/10.5771/0943-7444-2009-4-240.

Lo Vecchio, Nicholas. 2026. “For a Sociolexicology of Queer Naming: On Problematizing the LGBTQ Lexicon in Historical Linguistics.” In Battlefield Linguistics: Queer, Trans, and Feminist Interventions in Linguistic and Discursive Change, edited by Scott Burnett and Francesca Vigo, 159–178. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783111351568-008.

Martínez-Ávila, Daniel, and Melodie J. Fox. 2025. “Intersectionality in Knowledge Organization Revisited.” Informatio 30, no. 1: e201. https://doi.org/10.35643/Info.30.1.6.

Mongeon, Philippe, and Adèle Paul-Hus. 2016. “The Journal Coverage of Web of Science and Scopus: A Comparative Analysis.” Scientometrics 106: 213–228. https://doi.org/10.1007/s11192-015-1765-5.

Moura, Jonathan Ribeiro Farias de. 2018. “Da morfologia ao discurso: o caso do sufixo –ismo para denominar práticas homossexuais.” Ribanceira, no. 15: 99–111. https://periodicos.uepa.br/index.php/ribanceira/article/view/2139.

Olson, Hope A. 2002. The Power to Name: Locating the Limits of Subject Representation in Libraries. Dordrecht: Kluwer Academic Publishers.

Pamfile, Dana, Céline Bourquin, Laurent Michaud, Sebastien Brovelli, Pascale Pécoud, and Friedrich Stiefel. 2025. “What Is in a Chosen Name? An Exploratory Study on the Renaming Experiences of Transgender People.” International Journal of Transgender Health 26, no. 3: 592–610. https://doi.org/10.1080/26895269.2023.2301318.

Pinho, Fábio Assis. 2010. “Aspectos éticos em representação do conhecimento em temáticas relativas à homossexualidade masculina: uma análise da precisão em

linguagens de indexação brasileiras." PhD dissertation, Universidade Estadual Paulista. https://repositorio.unesp.br/handle/11449/103379.

Santana, Hiller Soares. 2014. "Questões de gênero e sexualidade no currículo escolar." Master's thesis, Universidade do Estado do Rio de Janeiro. http://www.bdtd.uerj.br/handle/1/10638.

Thelwall, Mike, Tracey J. Devonport, Meiko Makita, Kate Russell, and Lois Ferguson. 2023. "Academic LGBTQ+ Terminology 1900–2021: Increasing Variety, Increasing Inclusivity?" Journal of Homosexuality 70, no. 11: 2514–2538. https://doi.org/10.1080/00918369.2022.2070446.