



## Article

# Knowledge Organization and the Technological Challenges of the Big Data Era

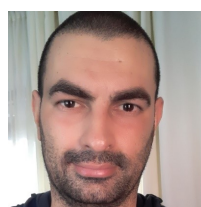
Marivalde Moacir Francelin<sup>1,\*</sup>, Fabio Orsi Meschini<sup>2</sup><sup>1</sup>Department of Information and Culture, School of Communications and Arts, University of São Paulo, 05508-220 São Paulo, Brazil<sup>2</sup>Central Library, Federal University of ABC, 09280-560 São Paulo, Brazil\*Correspondence: [marivalde@usp.br](mailto:marivalde@usp.br) (Marivalde Moacir Francelin)

Academic Editor: Natália Tognoli

Submitted: 11 December 2024   Revised: 4 March 2025   Accepted: 7 March 2025   Published: 27 April 2025



Marivalde Moacir Francelin holds a Ph.D. in Information Science from the University of São Paulo (USP). He is Associate Professor at the Department of Culture and Information of the University of São Paulo (USP). His main research interests include Information Science and Knowledge Organization.



PhD in Information Science (University of São Paulo - USP), Master in Information Science (São Paulo State University “Julio de Mesquita Filho” - Marília/SP), specialized in Public Management from the Federal University of Technology of Paraná, graduated in Library Science from the São Paulo State University “Julio de Mesquita Filho” - Marília/SP and Librarian at the Federal University of ABC - São Bernardo do Campo/SP. Areas of interest: Knowledge Organization, Big data, Artificial Intelligence, Information Metrics Studies and University Libraries.

## Abstract

Amid ongoing technological advances and their effects on the information cycle, this research examines how Knowledge Organization engages with Big Data technologies and analyzes the metatheoretical perspectives within this scientific domain. A bibliographic review of books and e-books, along with scientific articles from databases such as LISA, Scopus, Web of Science, and BRAPCI formed the theoretical framework. The empirical analysis focused on thematic content, theoretical influences (citations), and metatheoretical perspectives in the selected literature. The results show that technological developments significantly influence scientific practice in Information Science and Knowledge Organization. Nonetheless, their inherently social orientation fosters research that addresses both technological and societal concerns, providing insights into the Big Data context. The study concludes that managing large volumes of data today requires not only technological and semantic innovation but also a democratic approach—one best led by Information Science and Knowledge Organization in collaboration with disciplines such as Computer Science, Mathematics, and Statistics, underpinned by human and social guidelines.

**Keywords:** Knowledge Organization; Information Science; Big Data; technology

## 1. Introduction

Big Data is a resource-intensive phenomenon that inevitably raises ethical and privacy concerns. Addressing these issues is imperative, and recognizing Big Data as an inherently technosocial phenomenon is essential. Consequently, Big Data encompasses a broader scope beyond structured content.

Research endeavors that approach pressing societal issues, such as those seen in the context of Big Data and its challenges related to privacy, data access, and utilization, serve to underscore the significance of Information Science. Moreover, epistemological and methodological studies in this domain contribute to formulating organizational guidelines from a social perspective. Since its inception, the field

of Information Science has regarded technological devices as a means of managing the ever-increasing and evolving volume of information. The impact of technology on scientific practice in this field is evident. However, the social character of the field still requires greater emphasis in technological discussions. This social aspect of the field can inform debates on Big Data issues. As [Saracevic \(1996\)](#) and [Buckland \(2012\)](#) have noted, one role of Information Science is to facilitate the harmonization of humans, technologies, and information, prioritizing adapting technologies and information resources to human needs.

It is of the utmost importance to recognize that the exponential growth of data in virtual media is inextricably linked to the unrelenting advancement of technology,



resulting in an immense and complex volume of information that is challenging to comprehend and integrate fully. The critical issues brought about by Big Data are primarily within the fields of law, philosophy of science, and social sciences, particularly in identifying ethical, social, and economic questions. The data in question are not value-free; they carry theoretical biases and reflect human decisions, influencing their interpretation. The technological impact on the information cycle has led to the acquisition of various informational supports, indicating a growing need for organizational methods to manage the vast and diverse mass of data. Therefore, it is essential to emphasize areas such as Information Science and Knowledge Organization to address these organizational complexities.

In the context of this research and given the broad conceptual scope of the terms “Big Data” and “Technology”, the former is defined as “an information asset characterized by such a high volume, velocity, and variety that it requires specific technologies and analytical methods to transform it into value” (De Mauro et al, 2016, pp. 131). This endeavor aims to enhance the accessibility of this intricate data set to human cognition. Semeler and Pinto (2020) observed that the latter may encompass perspectives related to hardware, rules, and systems. In this context, it is defined as “a specific form of knowledge, encompassing unique ways of learning about and exploring the material world” (Semeler and Pinto, 2020, pp. 122). Consequently, it is associated with the technological tools and processes that drive the cycles of data flows.

In examining the context of Big Data, it is essential to acknowledge the ambivalence inherent in this data reality. On the one hand, Big Data has the potential to enhance and expand the Knowledge organization and representation. However, on the other hand, ethical complexities grow exponentially with this phenomenon, necessitating careful analysis. Consequently, the context of Big Data gives rise to a range of discussions pertaining to knowledge, including epistemological, methodological, ethical, and technological considerations.

This research investigates how Knowledge Organization comprehends and develops themes related to Big Data and analyzes the metatheoretical perspectives within this scientific production. In addition, the analysis includes metatheoretical perspectives associated with this body of scholarly work. This study employs a mixed-methods approach, combining a comprehensive literature review to establish the theoretical framework with a survey of books, e-books, and scientific articles from databases such as LISA, Scopus, Web of Science, and BRAPCI (a Brazilian digital platform for Information Science). In the empirical phase, a meticulous analysis was conducted to identify the central themes of the collected articles, the theoretical influences manifested through citations, and the metatheoretical perspectives reflected in the publications.

## 2. Knowledge Organization, Big Data, and the Digital Era

The dissemination of knowledge throughout society has historically constituted a considerable challenge. From the advent of the printing press in 1450, as pioneered by Gutenberg, to the encyclopedic validation of the 18th century, the definitions of knowledge for each historical moment were subject to recurrent change. The evolution of science and technology has had a profound impact on the ways in which people access data, information, and knowledge (Burke, 2000). This expansion of knowledge has given rise to the emergence of a publishing market for the production and distribution of knowledge. As a result, information has become a commodity, structured in various formats to facilitate access to knowledge across a range of fields, including technical and academic readings. This diversity in the presentation of information and knowledge has established a precedent that has been echoed in the advent of the Internet.

As Burke (2000) observed, the sheer quantity of information and knowledge available presents significant organizational and reliability challenges. The interconnection between different intellectuals is a necessary condition for comprehensive understanding of phenomena, given that it is impossible for any individual to grasp the entirety of available information. In his observations on the social history of knowledge, Burke identifies similarities with the reality of Big Data. These challenges pertain to the copious generation of data, the evolving nature of information formats and platforms, the prevalence of institutional monopolies in data management, and the trustworthiness of the available content. These elements persist and are addressed in the domains of data, information, and knowledge. From a comprehensive standpoint, the parallels between past and present in the dissemination of data, information, and knowledge suggest avenues for Knowledge Organization to enhance its resources in this vast data landscape.

The advent of Big Data has had a profound impact on scientific research, and consequently, on Knowledge Organization, which is confronted with the challenge of managing ever-increasing information volumes due to technological advancement. Hjørland (2016) proposes that the field of Knowledge Organization should be conceptualized “as a knowledge base that can be applied to all technological platforms”. Adopting a comprehensive view of Knowledge Organization allows for meaningful contributions to be made to a range of societal domains. The necessity for knowledge organization has become increasingly pressing, particularly in view of the advent of more dynamic services that align with technological reality.

The scope of Knowledge Organization’s scientific activities may give rise to the assumption that all objects and documents possess informational potential, thereby rendering it challenging to delineate the field’s boundaries. How-

ever, as [Gnoli \(2012\)](#) notes, the focus should be on the “content of the subject”. The material object is of secondary importance; what matters is its use in transmitting knowledge. The author underscores the potential for Big Data to facilitate new avenues of inquiry within the field of Knowledge Organization. It is essential to prioritize the flexible representation of information and knowledge, allowing for diverse combinations of the dimensions (phenomenon, perspective, and support) inherent to the subject.

A classification of information resources based on dimensions represented by phenomena, perspectives, and supports, as expressed by metadata, may prove an effective approach in the context of the vast quantities of data that are now a feature of the digital age. This would facilitate the formation of novel knowledge connections, which align with the expectations of Big Data contexts, such as data classification itself. The dimensions of classification, which encompass phenomenon, perspective, and support, would prove more conducive to comprehensive information retrieval. The incorporation of these elements into a variety of informational resources would facilitate access to dispersed knowledge within disciplinary classifications. An interdisciplinary approach would provide a vital alternative for understanding the technological scenario, contributing to the development of social, ethical, and organizational guidelines.

[San Segundo Manuel and Martínez-Ávila \(2014\)](#) posit that “digital thinking shapes our reality and its organizational form”. Therefore, novel structuring and organization forms are essential in the context of the vast data landscape. They posit that technological devices are integral to human cognition, emphasizing the urgency of discourse on technology’s influences. In its digital domain, Knowledge Organization pertains to both the materiality of technology itself and the cultural deposits within digital media.

[Guimarães et al \(2017\)](#) notes, the Knowledge Organization has emerged as a significant area of interest within interdisciplinary discussions within the field of Information Science. The theoretical and methodological foundations of this field are informed by the mediation of socially produced knowledge. In accordance with the conclusions presented by [Santos et al \(2019\)](#), the Knowledge Organization is of paramount importance within the domain of Information Science research. It is directly related to the processes of organization and representation of information and knowledge, which encompass technical, practical, and theoretical aspects. It can be asserted that research in this domain is of a strategic nature, addressing the challenges associated with diverse information processes. This is achieved by taking into account the inherent complexity of the relationship between information, users, and the conditions of information mediation ([Santos et al, 2019](#)).

As proposed by [San Segundo Manuel and Martínez-Ávila \(2014\)](#), advancements in technology exert a direct influence on the field of Knowledge Organization. The ad-

vent of digital network connections has the potential to facilitate the formation of a global network of interconnected individual intelligences, collectively capable of generating a significantly enhanced form of intelligence. The creation of organizational resources for the representation and semantic expansion of large data sets necessitates the development of innovative models, tools, and paradigms to effectively navigate the digital landscape. The rapid advancement of technology also gives rise to complexities inherent in the emulation of existing social relations and their associated inequalities in digital media. This renders the technological context a reflection of societal structures.

[San Segundo Manuel \(2013\)](#) underscores that the advent of technology has precipitated a paradigm shift in the way information is disseminated. This has given rise to novel social connections and unprecedented modes of navigating the informational and technological landscape. This paradigm is based on the compilation of large volumes of information to provide personalized products through algorithms, with the aim of making them more appealing to the target audience. The representation of the world is moving towards virtuality, with the result that the boundaries between the real and virtual are becoming increasingly blurred. This has the effect of amplifying ethical information conflicts. The virtualization of the world around us, centered on a few companies, can interfere with the exercise of citizenship.

The contemporary technological discourse has highlighted the pivotal role of the social paradigm in Knowledge Organization and Information Science, conceptualizing information as a phenomenon that incorporates social and human perspectives. The social significance of organizing, representing, and circulating informational and knowledge records leads to the understanding that their distribution and retrieval are regarded as social, political, and economic issues ([Santos et al, 2019](#)). In the context of massive data, these considerations become even more relevant and complex, given the ethical conflicts related to data use and access across various sectors of society.

[Martínez-Ávila \(2015\)](#) identifies a discrepancy between the theoretical and foundational aspects of Knowledge Organization and the most advanced technological applications in the field. This discrepancy may be attributed to the separation of theoretical studies from practical applications, which has resulted in a technological profile that is more applied in nature. The author proposes that the growing technological complexity has resulted in the marginalization of human experiences in Knowledge Organization, which remain fundamental for technological performance and research aligned with societal demands.

Theories and specialties in Knowledge Organization are of significant importance in their relationship with technology, given the superiority of human experiences and intelligences compared to artificial intelligence. However, [Martínez-Ávila \(2015\)](#) identifies several enhance-

ments needed in the relationship between Knowledge Organization and technology:

- Term assignment: Enhancing precision in classification and automatic indexing.
- Vocabulary alignments: Addressing incompatibility and lack of interoperability among domain vocabularies, widening the gap between machines and humans.
- System evaluation and projects: Improving communication with system developers for better usability.
- Domain analysis: Overcoming the predominance of individual bias in Cognitive Sciences and Artificial Intelligence research, while considering cultural, historical, and social aspects.
- Social tagging: Improving the quality of metadata attribution by laypersons or subject matter experts without Knowledge Organization knowledge.
- Sociological aspects and ethical commitments: Focusing on ethical, cultural, and social issues to reduce problems related to commercial interests.

In order to analyze the context of Big Data, it is necessary to consider the ambivalence inherent in the reality of the data, which presents a duality of potential outcomes. On the one hand, there is the potential for Knowledge Organization to improve and expand; on the other hand, however, the ethical complexities involved grow exponentially. The advent of big data has given rise to a plethora of discussions pertaining to epistemology, methodology, ethics, and technology. These discussions have a profound impact on a multitude of domains, including, but not limited to, privacy, intellectual property, politics, and society. The pivotal issue of data quality, provenance, credibility, and accuracy arises (Hajibayova and Salaba, 2018). It is of paramount importance to advocate for a meticulous examination of the impact of Big Data on Knowledge Organization. It is essential to strike a balance between the advantages of Big Data and the research on its social, ethical, and scientific implications. This will ensure the construction of Knowledge Organization systems that prioritize human development.

The introduction of automated processes in information indexing and retrieval has the effect of introducing a certain degree of complexity in the evaluation of retrieved items, particularly with regard to the precision of the subject matter. As Hjørland (2000) proposes, the objective of Knowledge Organization should be to optimize the representation and retrieval of information and knowledge for users. Automated processes rely on human interpretations and evaluations, which cannot be entirely supplanted. Consequently, the field of Knowledge Organization must develop methodologies related to domain analysis, which encompasses the study of knowledge domains and discourse communities. These are social groups that develop a common body of knowledge based on shared professional language, communication channels, and databases (Hjørland, 2000). The aforementioned domain studies bring the field closer to a sociological focus, prompting reflection on the

objectives, requirements, and aspirations that knowledge organization should serve. While commercial interests, as evidenced by major technology companies and social networks, are prioritized, social, democratic, and cultural principles must assume a prominent role in information matters.

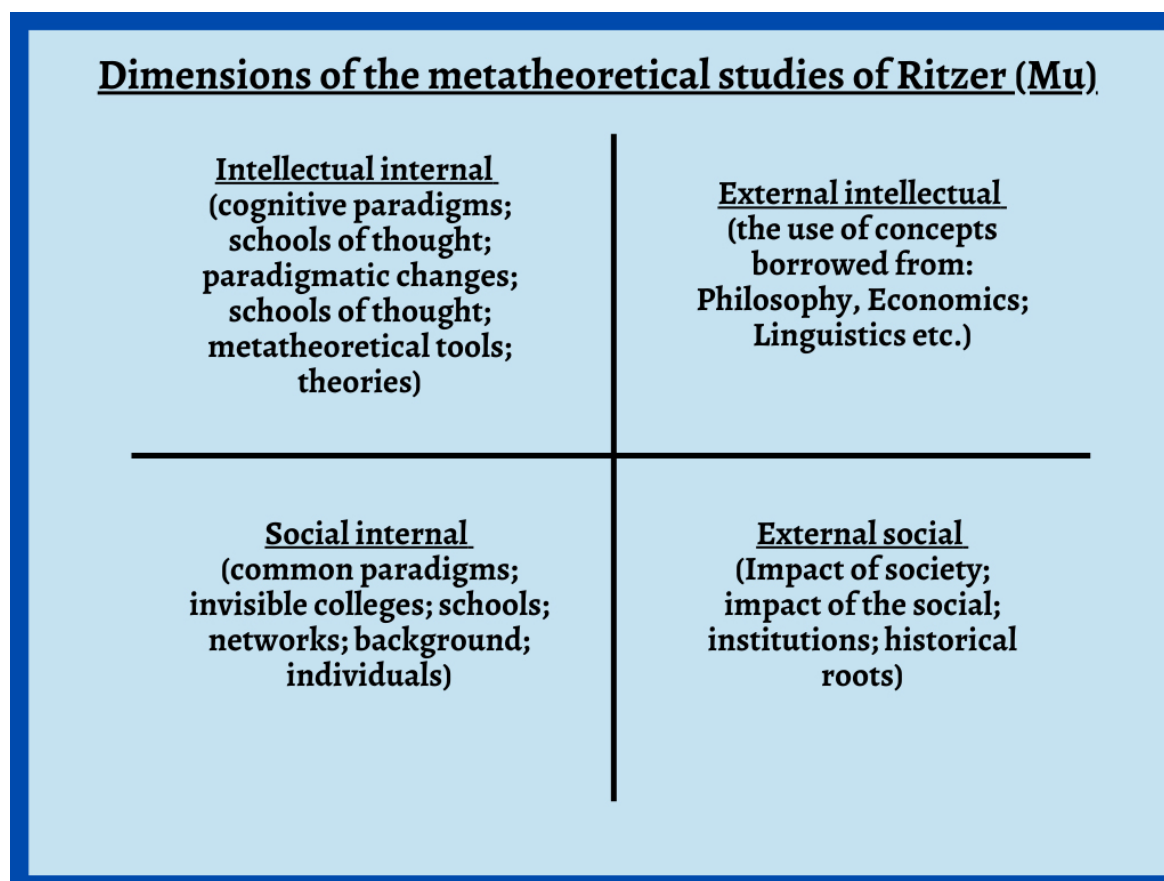
As observed by Khan et al (2017), the aggregation and correlation of data from disparate sources inherent to Big Data render analyses inherently complex. While data analysis is crucial for more accurate predictions, the primary challenge lies in the complexity of relating and analyzing these data relationships. While data security and user privacy are important considerations, the advent of Big Data presents an opportunity to extract relevant information for societal benefit.

Barité (2014) emphasizes the necessity of integrating controlled vocabularies with emerging technological resources. Such resources can facilitate data access through an organized presentation of controlled vocabularies, thereby meeting the demands of the user community. It is of paramount importance to acknowledge the relation between natural language and controlled vocabularies for optimal results, particularly given that data frequently concerns individuals. Ontologies and folksonomies highlight the substantial social dimension of Big Data. The adaptability of controlled vocabularies in the digital era is evident, suggesting that combining natural and controlled language, and considering users in the retrieval process, is a viable approach.

Szostak et al (2016) posit that the instabilities inherent to technological development should be regarded as opportunities for innovation and transformation in Knowledge Organization processes. In accordance with technological advancements and an interdisciplinary perspective, data connections are expanded, thereby maximizing informational value in a chaotic digital environment. In order to meet contemporary demands, including those related to digital databases and the development of the Semantic Web, current classification systems must adapt. The proposed metatheoretical classification aims to structure academic understandings and facilitate communication and synthesis in research on Big Data and Knowledge Organization. These contributions stimulate communication processes within and between specific domains.

Morin (1990) puts forth the proposition that the advancement of science and technology are inextricably linked. The relationship between science, technology, and knowledge is a complex one, comprising a multitude of intricate interactions and interdependencies. The advancement of science enables the production of technology, which in turn further advances science. It is therefore crucial to examine the interconnections between these three domains, as technology has the potential to expand the boundaries of knowledge. However, the creation of an excess of information through technology can also result in a lack of knowledge. In order to assimilate information and





**Fig. 1. Dimensions of Ritzer’s metatheoretical studies (Mu).** Fonte: Tognoli (2014).

construct knowledge, a theoretical framework is required that provides meaning and mental structures to support the construction of knowledge. Information overload can impede these processes by compromising cognitive connections and reflections, thus limiting knowledge.

The advancement of knowledge should encompass a synthesis of specialized domains, eschewing fragmentation and safeguarding the wisdom-knowledge that enriches our lives and facilitates advancement (Morin, 1990). It is imperative to consider the ramifications of technology on society, particularly the logic of artificial machines that pervade social life. These influences shape our conceptualizations of society, life, and humanity.

It is imperative to consider the impact of technology on society. By integrating ethical considerations with the potential benefits of scientific advancement, it is possible to foster a more humane social order. However, uncritical adherence to technological principles can be dangerous. To engage with knowledge, it is essential to engage in reflection based on substantial information, an awareness of one’s informational heritage, and an accurate interpretation of data, while avoiding a totalizing view that rationalizes all phenomena. This pursuit of rational unity can result in individuals being objectified for economic interests, rendering them susceptible to social manipulation.

### 3. Methodological Procedures

The research is of a bibliographic nature. It entailed a search of books and e-books, as well as the collection of scientific articles from the LISA, Scopus, Web of Science, and BRAPCI databases. The search terms employed included “Big Data”, “algorit\*”, “epistemolog\*”, “Information Science”, “Knowledge Organization”, “Information Organization”, “Ciência da Informação”, “Organização do Conhecimento”, and “Organização da Informação”.

The empirical component of the study entailed a thematic and metatheoretical analysis of articles sourced from the Scopus, Web of Science, LISA, and BRAPCI databases. Queries such as “Big Data” AND “Knowledge Organization” OR “Organização do Conhecimento” and “Big Data” AND “Information Organization” OR “Organização da Informação” were employed in the searches, which were conducted without temporal limitations to ensure comprehensive temporal retrieval. Filters were applied to restrict results to Information Science articles. A spreadsheet software and a word processing program were utilized to analyze the overlap between articles from different databases.

Subsequently, a comprehensive examination of the complete texts of the articles was undertaken to ascertain their alignment with the specified search terms, thereby eliminating inconsistencies. The initial search yielded 47

articles, which were then subjected to a process of elimination to identify those that were most pertinent to the topic of Big Data. This resulted in the exclusion of four articles that did not directly address the subject matter. The remaining 26 articles were then selected for further analysis, with nine being identified as particularly relevant for the development of technological resources to facilitate the management of Big Data. This technological context encompassed an in-depth examination of themes, theoretical influences (citations), and metatheoretical perspectives.

As Glänzel (2003) observes, citations serve to illustrate the prevailing paradigms within scientific communities, with a particular emphasis on methodological procedures and the identification of key scientists and publications. As Smiraglia (2011) asserts, citations assist in delineating the boundaries of a domain, with frequently cited authors representing the vanguard of research. Metatheory plays a pivotal role in the theoretical and methodological advancement of a subject, facilitating evolution through relationships, analyses, discussions, and reflections (Cas-tanha and Grácio, 2014). Metatheoretical and bibliometric studies are mutually reinforcing, offering a more comprehensive analytical framework that incorporates epistemological, sociological, and historical elements. This integrative approach allows for a nuanced understanding of an object from multiple perspectives, complementing quantitative and qualitative analyses.

In the context of this research, it is acknowledged that Ritzer's metatheoretical study model, designated as "Mu" (Fig. 1), is more appropriate for a comprehensive understanding of the Big Data context within Knowledge Organization. This model facilitates a profound comprehension of a theory, enabling a more expansive understanding through the following structure:

Up next, thematic analyses related to the technological context are presented.

#### 4. Technological Context and Its Themes

Fig. 2 offers a succinct overview of the primary research topics and authors within the technological context of Knowledge Organization.

The contributions from the research included in this context are presented in the following paragraphs:

- Salvador Benítez and Sánchez-Vigil (2020) acknowledge Big Data's impact on information connectivity and social behaviors. They draw parallels between the publishing industry and companies like Google and Netflix regarding Big Data analysis. They present software solutions for predicting reading preferences, seeing Big Data as an opportunity to provide better services. However, they briefly address the complex issues of massive data in their conclusion, focusing on information extraction as its wealth.

- Victorino et al (2018) propose a conceptual model for organizing Brazilian government data using ontologies.

They advocate for collaboration between Information Science and Computer Science to manage data mass and enhance government transparency. They briefly mention the need for data security and user privacy, seeing Big Data as a chance to extract relevant information for society.

- Du et al (2017) explores improving research management through researcher social networks influenced by scientific Big Data. They aim for information interoperability among research communities. By integrating different Knowledge Organization Systems, they argue that research social networks can handle community diversity within scientific Big Data.

- Ma et al (2017) researched enhancing information retrieval for users in a Big Data environment, affirming contributions to this phenomenon. They briefly mention user privacy without extensive discussion, proposing a conceptual model using Semantic Web techniques for addressing user needs in Big Data. They highlight user privacy's importance but do not delve deeply into its ethical and social challenges.

- Xia and Zhang (2016) present ways for libraries to use technologies to contribute to Big Data through linked data. They advocate for Digital Humanities practices to tap into Big Data's potential. Examples from Genealogy show comprehensive data analysis through linkages. They support Linked Data for reorganizing data and information to discover and construct new knowledge.

- Shiri (2014) emphasizes Knowledge Organization's role in Big Data through linked data and controlled vocabulary integration. He advocates Knowledge Organization as relevant for addressing Big Data's semantic issues. He underscores integrating controlled vocabularies with linked data to extract semantic content. The research highlights technological adaptation of vocabularies as a guiding principle for Knowledge Organization in the massive data context.

- Li et al (2019) emphasize the importance of technologies in digital library services, highlighting Information Science's relevance in Big Data. They cover Internet Plus, Artificial Intelligence, and Blockchain as resources for managing Big Data, seeing it as containing hidden knowledge to be unveiled through digital libraries. They emphasize the significance of digital libraries in handling vast data, albeit briefly mentioning human respect.

- Araújo and Lima (2019) highlight Information Science's potential in managing government data, which can help citizens understand information and assert their rights, establishing democratic guarantees. They see the Semantic Web as a way to handle Big Data, understanding that more data access expands democracy, ensuring social, economic, and informational rights.

- Sant'ana (2016) considers citizens' rights in the data lifecycle, recognizing Information Science's role in data issues with Mathematics and Computer Science support. The theoretical paper proposes guidelines for the data lifecycle.

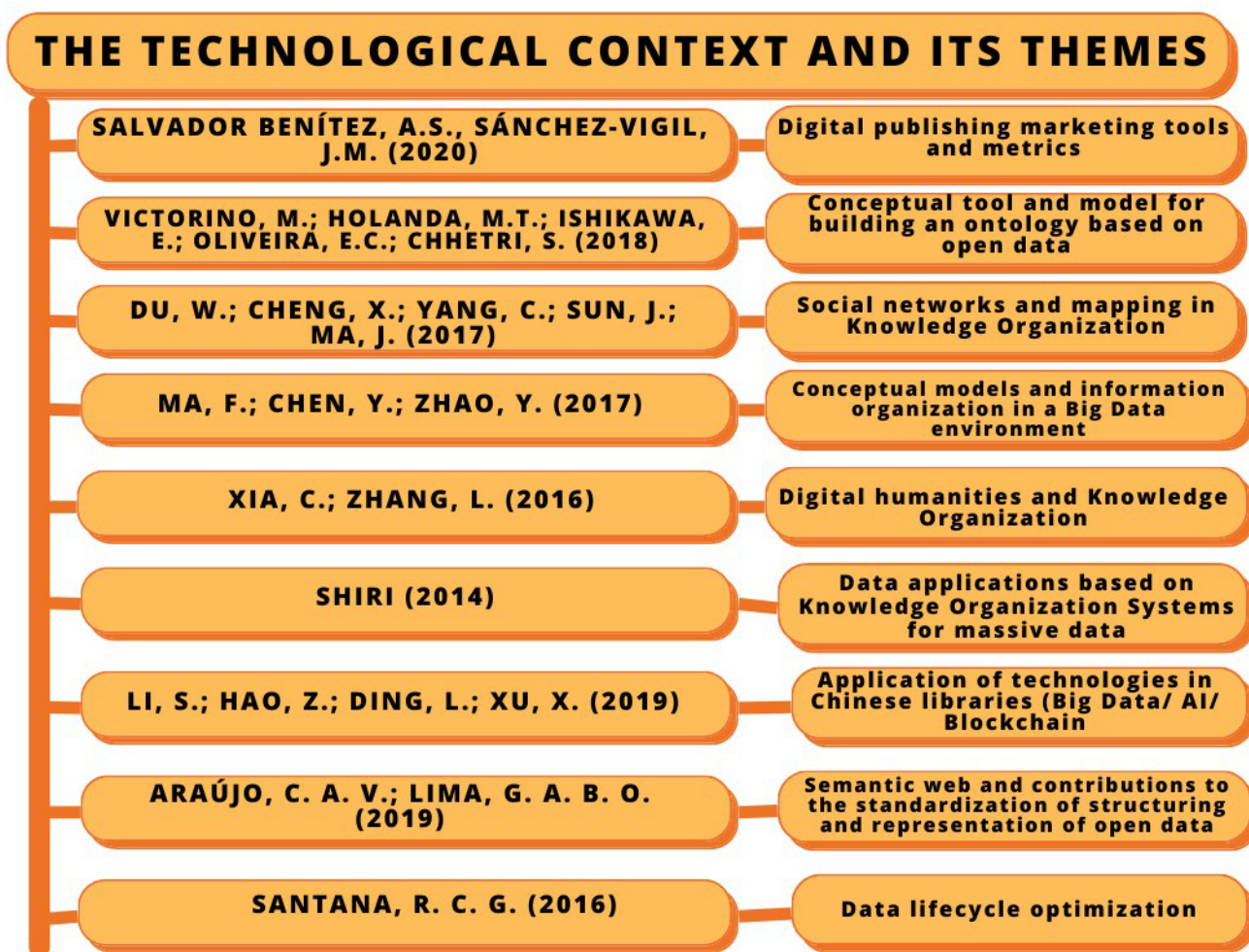


Fig. 2. The technological context and its themes.

cle, respecting privacy and rights, and emphasizes that data access expands informational possibilities, enabling democratic guarantees.

Research in the technological context is concerned with the enhancement of informational possibilities within the domain of Big Data, which in turn gives rise to new studies in the field of Knowledge Organization. While ethical and social discussions may not be extensively examined, the insights into the impact of Big Data on social behaviors and the enhancement of services for citizens through the extraction of pertinent information are noteworthy. Concerns about the retrievability of information and the construction of new knowledge through semantic enhancements play a pivotal role in these studies. Technological and semantic advancements in the management of vast data volumes represent a democratic assurance, led by Information Science, Knowledge Organization, and interdisciplinary collaborations with Computer Science, Mathematics, and Statistics.

## 5. Influences and Metatheoretical Characteristics of the Technological Context

Fig. 3 illustrates the theoretical influences and metatheoretical characteristics in the technological context.

It also demonstrates that references contain numerous authors, with 25 different authors mentioned on only two occasions. This suggests a comprehensive search for theoretical foundations due to the complexities of research involving technological, ethical, and social challenges. The primary theoretical influences are:

- [Berners-Lee et al \(2001\)](#), and [Berners-Lee \(2006\)](#): Discusses the Semantic Web's potential to evolve human knowledge by integrating concepts into a universal network, allowing data relationships and creating a web of linked data.
- [Bizer et al \(2007\)](#), [Bizer et al \(2011\)](#), and [Bizer and Cyganiak \(2012\)](#): Covers the Web of Data, linking and interconnecting data globally, describing and identifying digital resources, mapping, and linking database content on the Semantic Web, and addressing Big Data challenges.



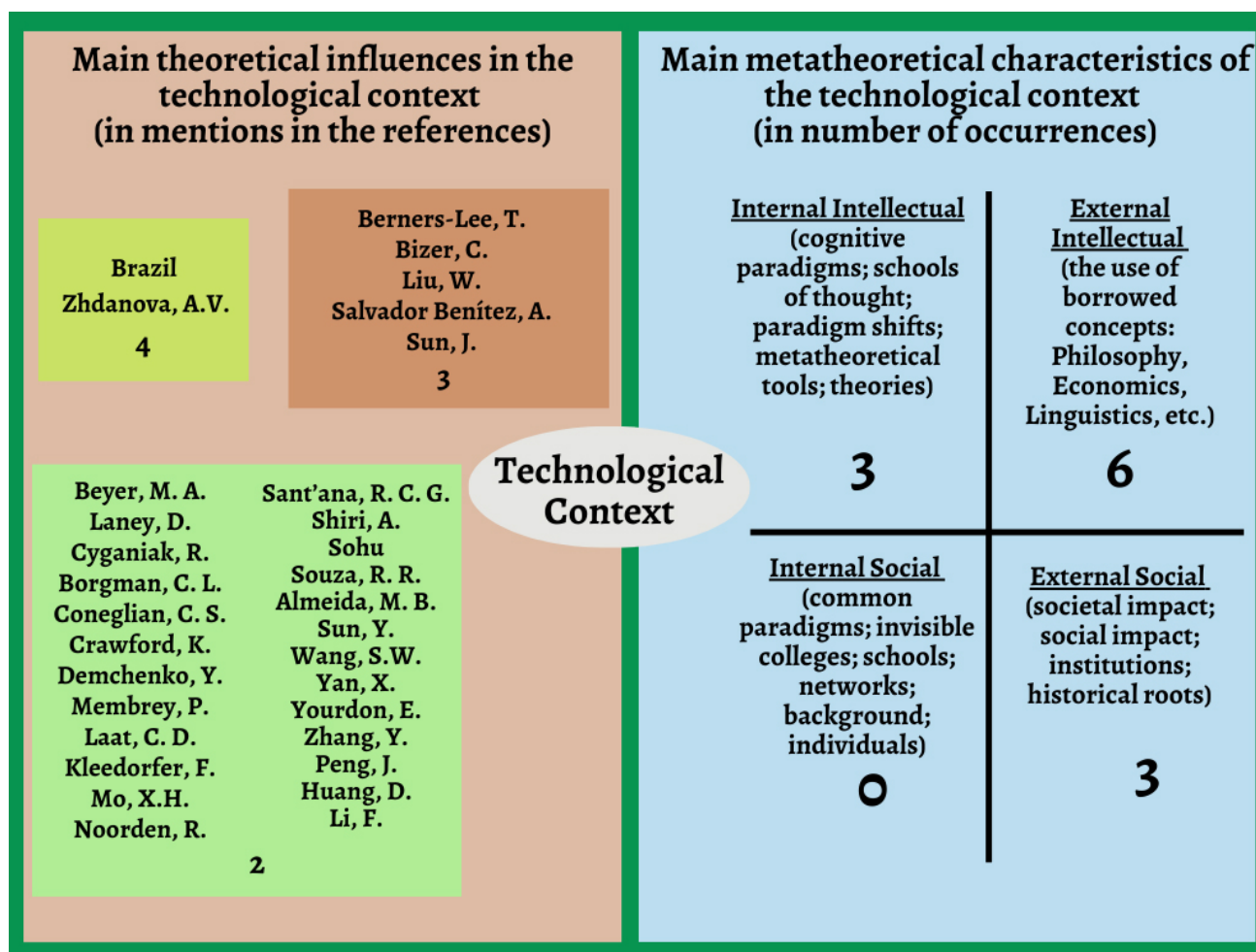


Fig. 3. Key theoretical influences and metatheoretical characteristics of the technological context.

- [Liu and Zhou \(2015\)](#): Discusses Machine Learning applications in libraries, analyzing technological impacts on Information Science, leading to adaptations and innovations.

- [Salvador Benítez and Gutierrez-David \(2010\)](#), [Salvador Benítez et al \(2014\)](#), and [Salvador Benítez \(2019\)](#): Examines economic advancements facilitated by data, user privacy, and data policies implemented by social networks due to the significant number of data-sharing users and their ethical implications.

- [Sun et al \(2014\)](#), and [Sun et al \(2015\)](#): Covers personalized scientific content recommendations through social research networks, enhancing researchers' assimilation of information using social computing and Big Data.

- [Brasil \(2011, 2014, 2016a,b\)](#): Discusses decrees and laws related to open data policies, digital governance, public information access, and internet usage in Brazil.

- [Zhdanova \(2005\)](#), [Zhdanova et al \(2005\)](#), [Zhdanova and Shvaiko \(2006\)](#), and [Zhdanova \(2008\)](#): Covers the development of community-oriented ontologies, more collaborative and aligned with user demands, utilizing the Semantic Web.

Metatheoretical perspectives in the technological context include:

- **External Intellectual**: [Salvador Benítez and Sánchez-Vigil \(2020\)](#); [Victorino et al \(2018\)](#); [Du et al \(2017\)](#); [Xia and Zhang \(2016\)](#); [Li et al \(2019\)](#); and [Sant'ana \(2016\)](#). These researchers address the use of software from Computer Science to enhance editorial services, interdisciplinary collaboration for government transparency, information interoperability among researchers, Digital Humanities practices to handle Big Data, Internet Plus, AI, and Blockchain for libraries, and Information Science's contributions to Big Data with Mathematics and Computer Science support.

- **Internal Intellectual**: [Ma et al \(2017\)](#); [Shiri \(2014\)](#); and [Araújo and Lima \(2019\)](#). These researchers focus on improving information retrieval using the Semantic Web, integrating controlled vocabularies with linked data for semantic development, and creating tools to enhance citizens' understanding of government data through the Semantic Web.

- **External Social**: [Du et al \(2017\)](#); [Araújo and Lima \(2019\)](#); and [Sant'ana \(2016\)](#). These researchers highlight enhancing research management through social networks,



information interoperability, conscious use of government data as a democratic guarantee, and including democratic and social issues in the data lifecycle.

In summary, the theoretical foundations of technology-related research are derived from various researchers due to the extensive and intricate nature of this field. The authors of the references address a number of issues, including the Semantic Web, Linked Data, the Data Web, database interconnections, machine learning, data policies in social networks, scientific social networks, content personalization, social computing, open data, digital governance, and community-driven ontology. These issues highlight the efforts of Information Science and Knowledge Organization to integrate technological tools for managing the vast amounts of data that are characteristic of the present era. They also touch on discussions of privacy, democracy, and citizenship, which are important social aspects that warrant further exploration.

Regarding metatheoretical characteristics:

- **External Intellectual:** There is a strong search for theoretical elements and technological resources from various fields to expand data semantics and increase informational possibilities for users. This involves collaboration and interdisciplinarity from Computer Science, Digital Humanities, AI, Blockchain, and Mathematics, focusing on interoperability between scientific domains.

- **Internal Intellectual:** This aims to develop existing resources, innovating and adapting to Big Data. It addresses issues like Information Retrieval, Controlled Vocabularies, and the Semantic Web, with a developmental focus.

- **External Social:** This aspect highlights the social relevance in constructing theories and methodologies in Knowledge Organization. Themes include the social impacts of extensive data use in scientific work and conscious use of government data to guarantee democratic rights while respecting user privacy. Social and democratic issues should be key focuses for Big Data research in Knowledge Organization.

## 6. Conclusions

The field of Knowledge Organization employs analytical and informational representation methods to address the complexities inherent in Big Data, thereby avoiding confusion between data acquisition and interpretation (Borgman, 2015). The field facilitates the acquisition of data and its subsequent interpretation, thereby contributing to semantic expansion, particularly in the context of scientific data.

Data analysis in a Big Data context is of significant importance for connecting data, simplifying search parameters, and making accurate decisions by accessing diverse sources and perspectives. However, evaluating useful data within this abundance remains a complex challenge.

Knowledge Organization should prioritize the user, integrating diversity and interdisciplinary approaches into

current information issues. The social impact of its activities redefines the centrality of the user, as without users, there is no Information Science and Knowledge Organization. Adaptability in the digital age is essential for the field's development, leading to a reevaluation of information boundaries and organizational procedures in a broad and discontinuous flow.

The data phenomenon represents a novel paradigm for Information Science and Knowledge Organization, exerting a profound and far-reaching impact on information activities. The effective management of the data abundance and its complex algorithms hinges on a thorough understanding of the human factors involved. Technological advancements should enhance users' quality of life, with human intervention serving as a guiding force in the development of algorithmic models. A social focus can facilitate the establishment of ethical parameters, with interdisciplinary relationships contributing to the ethical dimension through studies on information ethics and social perspectives.

Information Science and Knowledge Organization are vital in discovering knowledge within Big Data. Technological solutions should address everyday social problems and contribute to more efficient societal management. It is imperative that research strategies leverage Big Data for organizational development, emphasizing Knowledge Organization's social function and addressing ethical issues in commercial data management.

This social perspective confirms the necessity to analyze and understand Big Data through ethical and technological lenses, considering the human factor's complexities. It is essential to engage in conceptual and theoretical reflections to address the technological aspects of Big Data in Knowledge Organization. Epistemological contributions are necessary to reduce semantic ambiguities, ensuring access and use of data for the promotion of informational and social development.

In order to contribute to this social perspective on technological issues, it is proposed that the following thesis of Big Data be adopted: a comprehensive data phenomenon requiring an infotechnological structure based on technical, social, and ethical guidelines. It is a technosocial phenomenon that requires resources to address ethical and privacy issues. Interdisciplinary approaches promote the integration of traditional information systems with innovative technological perspectives. The actions of Knowledge Organization should focus on expanding informational possibilities and knowledge assimilation by individuals in various environments, physical or virtual.

## Availability of Data and Materials

All data points generated or analyzed during this study are included in this article and there are no further underlying data necessary to reproduce the results.

## Author Contributions

FOM designed and collected the survey data. FOM and MMF analyzed and discussed the results. MMF prepared the final version of the article. Both authors contributed to editorial changes in the manuscript. Both authors read and approved the final manuscript. Both authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

## Acknowledgment

Not applicable.

## Funding

This research received no external funding.

## Conflict of Interest

The authors declare no conflict of interest.

## References

- Araújo CAV, Lima GABO. 2019. Structuring and semantic representation of big data in the context of government databases. *Múltiplos Olhares em Ciência da Informação*. 2019; 9. (In Portuguese)
- Barité M. Vocabulary control in the digital age: a conceptual review. *Scire*. 2014; 20: 99–108. (In Spanish)
- Berners-Lee T. Linked data-design issues. 2006. Available at: <http://www.w3.org/DesignIssues/LinkedData.html> (Accessed: 3 November 2024)
- Berners-Lee T, Hendler J, Lassila O. The semantic web. *Scientific American*. 2001; 285: 29–37.
- Bizer C, Boncz PA, Brodie ML, Erling O. The meaningful use of Big Data: four perspectives-four challenges. *Sigmod Record*. 2011; 40: 56–60. <https://doi.org/10.1145/2094114.209412>
- Bizer C, Cyganiak R. D2R server-publishing relational databases on the semantic web. In Poster at the 5th international semantic web conference. 2012.
- Bizer C, Cyganiak R, Gauß T. The RDF book mashup: from web APIs to a web of data. *CEUR-WS*. 2007; 248: 1–6.
- Borgman C. Big data, little data, no data: scholarship in the networked world. MIT Press: Cambridge. 2015. <https://doi.org/10.7551/mitpress/9963.001.0001>
- Brasil. Decreto no. 8.638. Diário Oficial da União (January). 2016a. Available at: <http://www2.camara.leg.br/legin/fed/decret/2016/decreto-8638-15-janeiro-2016-782270-publicacaooriginal-149245-pe.html> (Accessed: 3 November 2024) (In Portuguese)
- Brasil. Decreto no. 8.777. Diário Oficial da União (May). 2016b. Available at: <http://www2.camara.leg.br/legin/fed/decret/2016/decreto-8777-11-maio-2016-783082-norma-pe.html> (Accessed: 3 November 2024) (In Portuguese)
- Brasil. Lei no. 12.527. Diário Oficial da União (November). 2011. Available at: [https://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2011/lei/112527.htm](https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/112527.htm) (Accessed: 3 November 2024) (In Portuguese)
- Brasil. Lei no. 12.965: estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil. Diário Oficial da União (April). 2014. Available at: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2014/lei/112965.htm#art32](http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/112965.htm#art32) (Accessed: 3 November 2024) (In Portuguese)
- Buckland M. What kind of science can information science be? *Journal of the American Society for Information Science and Technology*. 2012; 63: 1–7. <https://doi.org/10.1002/asi.21656>
- Burke P. A Social History of Knowledge: From Gutenberg to Doderot. Polity Press: Cambridge. 2000.
- Castanha RCG, Grácio MCC. Bibliometrics contribution to the metatheoretical and domain analysis studies. *Knowledge Organization*. 2014; 4: 171–174. <https://doi.org/10.5771/0943-7444-2014-2-171>
- De Mauro A, Greco M, Grimaldi M. A formal definition of Big Data based on its essential features. *Library Review*. 2016; 65: 122–135. <https://doi.org/10.1108/LR-06-2015-0061>
- Du W, Cheng X, Yang C, Sun J, Ma J. Establishing interoperability among knowledge organization systems for research management: a social network approach. *Scientometrics*. 2017; 112: 1489–1506. <https://doi.org/10.1007/s11192-017-2457-0>
- Glänzel W. Bibliometrics as a research field: a course on theory and application of bibliometric indicators. s.n: Bélgica. 2003.
- Gnoli C. Metadata about what? Distinguishing between ontic, epistemic, and documental dimensions in knowledge organization. *Knowledge Organization*. 2012; 39: 268–275. <https://doi.org/10.5771/0943-7444-2012-4-268>
- Guimarães JAC, Martinez-Ávila D, Cândido GG, Silva AP da, Piva LMR, Fioravanti DCB. The theoretical Dimension of Domain Analysis in the Brazilian Scientific Production in Information Science. In Pinho FA, Guimarães JAC (eds.) *Memória, tecnologia e cultura na organização do conhecimento*. UFPE: Recife. 2017.
- Hajibayova L, Salaba A. Critical questions for big data approach in knowledge representation and organization. In *Challenges and opportunities for Knowledge Organization in the Digital Age* (pp. 144–151). Ergon: Baden-Baden. 2018.
- Hjørland B. Should librarians organize all information on the Internet? *Human IT*. 2000; 4: 48–70.
- Hjørland B. Knowledge organization (KO). *Knowledge Organization*. 2016; 43: 475–484. <https://doi.org/10.5771/0943-7444-2016-6-475>
- Khan S, Liu X, Shakil KA, Alam M. A survey on scholarly data: From big data perspective. *Information Processing & Management*. 2017; 53: 923–944. <https://doi.org/10.1016/j.ipm.2017.03.006>
- Li S, Hao Z, Ding L, Xu X. Research on the application of information technology of Big Data in Chinese digital library. *Library Management*. 2019; 40: 518–531. <https://doi.org/10.1108/LM-04-2019-0021>

- Liu W, Zhou D. Forward-looking of the next decade library technology trends. *Library Journal*. 2015; 34: 4–12. (In Chinese)
- Ma F, Chen Y, Zhao Y. Research on the organization of user needs information in the big data environment. *The Electronic Library*. 2017; 35: 36–49. <https://doi.org/10.1108/EL-07-2015-0130>
- Martínez-Ávila D. Knowledge Organization in the intersection with information technologies. *Knowledge Organization*. 2015; 42: 486–498. <https://doi.org/10.5771/0943-7444-2015-7-486>
- Morin E. *Science avec Conscience*. Fayard: Paris. 1990.
- Salvador Benítez A. Between privacy and business: terms of use and data policy in social networks. *Derecom*. 2019; 27: 1–23. (In Spanish)
- Salvador Benítez A, Gutierrez-David ME. Social networks and mass media: legal challenges. *Profesional de la Informacion*. 2010; 19: 667–674.
- Salvador Benítez A, Olivera Zaldúa M, Sánchez Vigil JM. Social network and media audiovisual material use policy. In *Tecnología y narrativa audiovisual* (pp. 717–735). Fragua: Madrid. 2014.
- Salvador Benítez A, Sánchez-Vigil JM. Tools and metrics in editorial marketing: from big data to artificial intelligence. *Scire*. 2020; 26: 35–46. <https://doi.org/10.54886/scire.v26i1.4682>
- San Segundo Manuel R. Panorama de investigación en Organización del conocimiento en su dimensión epistemológica. In Dodebei V, Guimarães JAC (eds.) *Complexidade e organização do conhecimento, desafios de nosso século*. ISKO-Brasil: Rio de Janeiro. 2013.
- San Segundo Manuel R, Martínez-Ávila D. Digital as a hegemonic medium for epistemology and knowledge organization. In Babik W (ed.) *Knowledge Organization in the 21st Century: Between Historical Patterns and Future Prospects*. Verlag: Baden-Baden. 2014.
- Sant’ana RCG. Data life cycle: a perspective from the Information Science. *Informação & Informação*. 2016; 21: 116–142. (In Portuguese)
- Santos RF dos, Neves DAB, Souza ED de. The Knowledge Organization as domain of study of Information Science: a reflection from the epistemological aspects. In Barros THB, Tognoli NB (eds.) *Organização do conhecimento responsável: promovendo sociedades democráticas e inclusivas*. UFPA: Belém. 2019. (In Portuguese)
- Saracevic T. Information Science: origin, evolution and relations. *Perspectivas em Ciência da Informação*. 1996; 1: 41–62. (In Portuguese)
- Semeler AR, Pinto AL. Librarianship in the Age of Data Science: Data Librarianship Venn Diagram. In Mugnaini R (ed.) *Data and Information in Online Environments*. DIONE 2020. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. Springer: Cham. 2020. [https://doi.org/10.1007/978-3-030-50072-6\\_10](https://doi.org/10.1007/978-3-030-50072-6_10)
- Shiri A. Linked data meets Big data: a knowledge organization systems perspective. *Advances in Classification Research Online*. 2014; 24: 16–20. <https://doi.org/10.7152/acro.v24i1.14672>
- Smiraglia RP. ISKO 11’s diverse bookshelf: an editorial. *Knowledge Organization*. 2011; 38: 179–189. <https://doi.org/10.5771/0943-7444-2011-3-179>
- Sun J, Ma J, Liu Z, Miao Y. Leveraging content and connections for scientific article recommendation in social computing contexts. *The Computer Journal*. 2014; 57: 1331–1342. <https://doi.org/10.1093/comjnl/bxt086>
- Sun J, Xu W, Ma J, Sun J. Leverage RAF to find domain experts on research social network services: A big data analytics methodology with MapReduce framework. *International Journal of Production Economics*. 2015; 165: 185–193. <https://doi.org/10.1016/j.ijpe.2014.12.038>
- Szostak R, Gnoli C, López-Huertas M. *Interdisciplinary Knowledge Organization*. Springer: Cham. 2016.
- Tognoli NB. A construção teórica da Diplomática: em busca da sistematização de seus marcos teóricos como subsídio aos estudos arquivísticos. *Cultura Acadêmica*: São Paulo. 2014.
- Victorino M, de Holanda MT, Ishikawa E, Oliveira EC, Chhetri S. Transforming open data to linked open data using ontologies for information organization in big data environments of the brazilian government: the brazilian database government open linked data–dbgoldbr. *Knowledge Organization*. 2018; 45: 443–466. <https://doi.org/10.5771/0943-7444-2018-6-443>
- Xia C, Zhang L. The application of linked data in Shanghai library’s service of genealogy digital humanities. *Library Journal*. 2016; 35: 26–34. (In Chinese)
- Zhdanova AV. Towards a community-driven ontology matching. In K-CAP ‘05: Proceedings of the 3rd international conference on Knowledge capture (pp. 221–222). 2005. <https://doi.org/10.1145/1088622.1088678>
- Zhdanova AV. Community-driven ontology construction in social networking portals. *Web Intelligence and Agent Systems: An International Journal*. 2008; 6: 93–121.
- Zhdanova AV, Krummenacher R, Henke J, Fensel D. Community-driven ontology management: DERI case study. The 2005 IEEE/WIC/ACM International Conference on Web Intelligence (WI’05) (pp. 73–79). 2005.
- Zhdanova AV, Shvaiko P. Community-driven ontology matching. In *European Semantic Web Conference* (pp. 34–49). Springer: Berlin. 2006. [https://doi.org/10.1007/11762256\\_6](https://doi.org/10.1007/11762256_6)