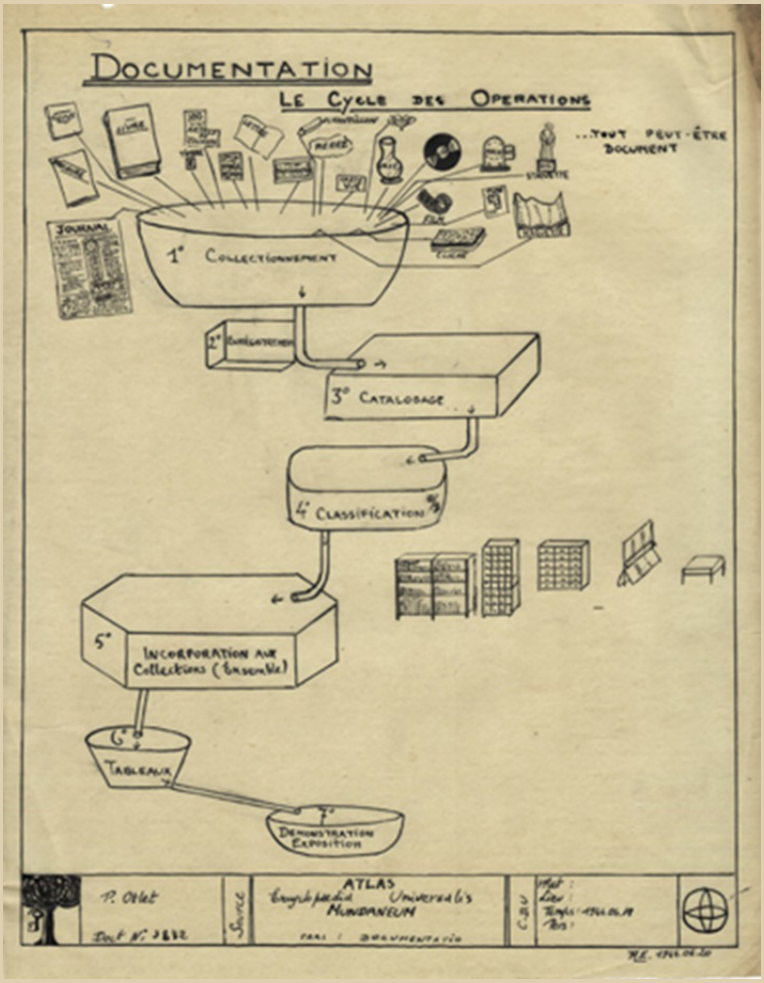
AIDAinformazioni

RIVISTA SEMESTRALE DI SCIENZE DELL'INFORMAZIONE

NUMERO 3-
ANNO 43
LUGLIO-DICEMBRE 2025



CACUCCI EDITORE
BARI

AIDAinformazioni

Rivista semestrale di Scienze dell'Informazione

Fondata nel 1983 da Paolo Bisogno

Proprietario della rivista:
Università della Calabria

Direttore Scientifico:
Roberto Guarasci, *Università della Calabria*

Direttore Responsabile:
Fabrizia Flavia Sernia

Comitato scientifico:
Anna Rovella, *Università della Calabria*;
Maria Guercio, *Sapienza Università di Roma*;
Giovanni Adamo, *Consiglio Nazionale delle Ricerche* †;
Claudio Gnoli, *Università degli Studi di Pavia*;
Ferruccio Diozzi, *Centro Italiano Ricerche Aerospaziali*;
Gino Roncaglia, *Università della Tuscia*;
Laurence Favier, *Université Charles-de-Gaulle Lille 3*;
Madjid Ihadjadene, *Université Vincennes-Saint-Dénis Paris 8*;
Maria Mirabelli, *Università della Calabria*;
Agustín Vivas Moreno, *Universidad de Extremadura*;
Douglas Tudhope, *University of South Wales*;
Christian Galinski, *International Information Centre for Terminology*;
Béatrice Daille, *Université de Nantes*;
Alexander Murzaku, *College of Saint Elizabeth*, USA;
Federico Valacchi, *Università di Macerata.*

Comitato di redazione:
Antonietta Folino, *Università della Calabria*;
Erika Pasceri, *Università della Calabria*;
Maria Taverniti, *Consiglio Nazionale delle Ricerche*;
Maria Teresa Chiaravalloti, *Consiglio Nazionale delle Ricerche*;
Assunta Caruso, *Università della Calabria*;
Claudia Lanza, *Università della Calabria.*

Segreteria di Redazione:
Valeria Rovella, *Università della Calabria*

Editrice: Cacucci Editore S.a.s.
Via D. Nicolai, 39 – 70122 Bari (BA)
www.cacucciedITore.it
e-mail: riviste@cacucciedITore.it
Telefono 080/5214220

AIDAinformazioni

Rivista semestrale di Scienze dell'Informazione

«AIDAinformazioni» è una rivista scientifica che pubblica articoli inerenti alle Scienze dell'Informazione, alla Documentazione, all'Archivistica, alla Gestione Documentale e all'Organizzazione della Conoscenza ma amplia i suoi confini in ulteriori campi di ricerca affini quali la Terminologia, la Linguistica Computazionale, la Statistica Testuale, ecc. È stata fondata nel 1983 quale rivista ufficiale dell'Associazione Italiana di Documentazione Avanzata e nel febbraio 2014 è stata acquisita dal Laboratorio di Documentazione dell'Università della Calabria. La rivista si propone di promuovere studi interdisciplinari oltre che la cooperazione e il dialogo tra profili professionali aventi competenze diverse, ma interdipendenti. I contributi pubblicati affrontano questioni teoriche, metodologie adottate e risultati ottenuti in attività di ricerca o progettuali, definizione di approcci metodologici originali e innovativi, analisi dello stato dell'arte, ecc.

«AIDAinformazioni» è riconosciuta dall'ANVUR come rivista di Classe A per l'Area 11 – Gruppo Scientifico Disciplinare 11/HIST-04 – Scienze del libro, del documento e storico-religiose e come rivista scientifica per le Aree 10 – Scienze dell'antichità, filologico-letterarie e storico-artistiche; 11 – Scienze storiche, filosofiche, pedagogiche e psicologiche; 12 – Scienze giuridiche; 14 – Scienze politiche e sociali. È anche annoverata dall'ARES (Agence d'évaluation de la recherche et de l'enseignement supérieur) tra le riviste scientifiche dell'ambito delle Scienze dell'Informazione e della Comunicazione. La rivista è, inoltre, indicizzata in: ACNP – Catalogo Italiano dei Periodici; BASE –Bielefeld Academic Search Engine; ERIH PLUS – European Reference Index for the Humanities and Social Sciences – EZB – Elektronische Zeitschriftenbibliothek – Universitätsbibliothek Regensburg; Gateway Bayern; KVK – Karlsruhe Virtual Catalog; The Library Catalog of Georgetown University; SBN – Italian union catalogue; Ulrich's; Union Catalog of Canada; LIBRIS – Union Catalogue of Swedish Libraries; Worldcat.

I contributi sono valutati seguendo il sistema del *double blind peer review*: gli articoli ricevuti sono inviati in forma anonima a due referee, selezionati sulla base della loro comprovata esperienza nei topics specifici del contributo in valutazione.

AIDAinformazioni

Anno 43

N. 3-4 – luglio-dicembre 2025

CACUCCI EDITORE

BARI

PROPRIETÀ LETTERARIA RISERVATA

© 2025 Cacucci Editore – Bari
Via Nicolai, 39 – 70122 Bari – Tel. 080/5214220
http://www.cacucciedi tore.it e-mail: info@cacucci.it

Ai sensi della legge sui diritti d'Autore e del codice civile è vietata la riproduzione di questo libro o di parte di esso con qualsiasi mezzo, elettronico, meccanico, per mezzo di fotocopie, microfilms, registrazioni o altro, senza il consenso dell'autore e dell'editore.

Sommario

Giovanna Aracri, Antonietta Folino, Maria Teresa Guaglianone, Marco Manna, Identification of thesaurus relationships with the support of ChatGPT and Gemini 9

Alessandra Bonsignori, Terminologia giuridica e organizzazione della conoscenza 51

Maria Teresa Chiaravalloti, Maria Teresa Guaglianone, Elena Cardillo, Metadatazione delle tipologie di documenti clinici: un approccio alla standardizzazione mediante LOINC 73

Carla Beatriz Marques Felipe, Carla Maria Martellote Viola, Raimunda Fernanda dos Santos, Domain Analysis on Data Management 91

Claudio Gnoli, Emanuela Valenzano, Che faccetta sei? 111

Manuela Grillo, Strumenti di convergenza tra archivi e biblioteche: i thesauri 121

Bianca La Manna, Rappresentare gli atti notarili medievali in Linked Open Data (LOD) 141

Claudia Lanza, L'univocità e la variabilità terminologica a supporto della rappresentatività delle risorse semantiche di organizzazione e rappresentazione della conoscenza specialistica 163

Larissa Moraes Martins, Carlos Eduardo da Silva Carvalho, Ana Cristina de Albuquerque, Mario Barité, Possibilities for using cultural warrant to complement literary work 181

Bianca Therezinha Carvalho Panisset, Lucia Maria Velloso de Oliveira, Leticia Souza da Costa Sampaio, Camilla Campoi de Sobral, Identification of types of documents and knowledge organization: transparency and access in records management 197

SoYoung Yoon, Claudio Gnoli, Ziyoung Park, Applying Integrative Levels Classification to Korean KOS Registry 211

Emanuela Nicole Donato, Grazia Serratore, Maria Taverniti, Maria Teresa Chiaravalloti, L'evoluzione della struttura documentale nelle cartelle cliniche psichiatriche dell'archivio storico del manicomio di Girifalco (1881-1978) 237

Lorenza Talarico, Manuscripts in the Digital Age 263

Maria Lígia Triques, Ananda Fernanda de Jesus, Mario Barité, José Eduardo Santarem Segundo, Ana Cristina de Albuquerque, Vocabularies in the digital context 275

Marcin Trzmielewski, Context-Driven Use of Automatic Tools in Medical Ontology Design: A Case Study in Allergology 291

Carla Maria Martellote Viola, Carla Beatriz Marques Felipe, Exploring the relationships between the Cross-Cutting Women's Agenda - Multi-Year Plan 2024-2027 and the Knowledge Organization of legislative proposals under processing on women in Brazil 313

Elena Cencetti, Camillo Carlo Pellizzari di San Girolamo, Elisabetta Viti, Modellizzazione delle conoscenze: Thesaurus Nuovo soggettario e Wikidata, due strumenti a confronto 325

Pontish Yeramyan, Alexander Murzaku, Nick Uva, Curt Anderson, The transferability of brilliance 353

Ana Flávia Dias Zammataro, Ana Cristina de Albuquerque, Giulia Crippa, Lucia Sardo, Carlo Bianchini, Maria Lígia Triques, ISKO Italy in perspective 375

Contributi

AIDAinformazioni
ISSN 1121–0095
ISBN 979-12-5965-639-1
DOI 10.57574/596563914
pag. 91-110 (luglio-dicembre 2025)
Open access article licensed under CC-BY-NC-SA

Domain Analysis on Data Management

A View from Brazilian Information Science

Carla Beatriz Marques Felipe*, Carla Maria Martellote Viola**,
Raimunda Fernanda dos Santos***

Abstract: The general objective of this work is to describe what constitutes the domain of Data Management in Brazilian Information Science. To this end, it develops a domain analysis on the subject. Two approaches were selected for development: bibliometric study and epistemology. The methodology used is bibliographic and documentary, with the analysis of Brazilian Postgraduate Programs in Information Science that offer disciplines on Data Management and bibliographic analysis on the subject. As a result, it is observed that what is researched in the Postgraduate Programs is reflected in the bibliographic production. In addition, it is noted that Brazilian Information Science develops more research-on-Research Data Management.

Keywords: Data Management, Domain Analysis, Brazilian Information Science.

1. Introduction

Due to e-Science and the Open Science Movement – which advocates for making information accessible to all – the sharing of research data has gained prominence in the academic context. In this case, data are considered products generated throughout scientific research, in various formats (texts, images, sounds, spreadsheets, videos, etc.) and across multiple fields of knowledge.

Sharing and reuse by researchers will only be possible if efficient data management is in place. Therefore, data management must be considered throughout the entire research process and should be developed through actions related to the organization of data information and the sharing infrastructure, always aiming for the data to remain usable even after the research has concluded.

* Federal University of Rio de Janeiro, Course in Library Science and Information Unit Management; Rio de Janeiro, Brasil. felipecarla12@gmail.com. https://orcid.org/0000-0001-5277-9165.

** Instituto Brasileiro de Informação em Ciência e Tecnologia; Brasília -DF, Brasil. viola.carla@gmail.com. https://orcid.org/0000-0002-0895-8163.

*** Federal University of Rio Grande do Norte, Library Science Course; Natal, Brasil. raimunda.fernanda@ufrn.br. https://orcid.org/0000-0002-7750-3269.

2025, The Author(s). This is an open access article, free of all copyright, that anyone can freely read, download, copy, distribute, print, search, or link to the full texts or use them for any other lawful purpose. This article is made available under a Creative Commons Attribution Share-Alike 4.0 International License. AIDAinformazioni is a journal of the Department of Culture, Education and Society, University of Calabria, Rende (CS), Italy, published by Cacucci Editore, Bari, Italy.

According to Alonso-Arévalo (2019, 82), data management operates along two lines: «long-term preservation of datasets in storage and sharing systems, and reuse of datasets for research purposes». As actions related to data preservation are implemented, sharing and subsequent reuse are facilitated.

The BRIET[1] research group has, as one of its goals, the development of studies on the concept of research data and the processes involved in transforming data into information and scientific knowledge. In addition, it conducts research on Knowledge Organization.

Knowledge Organization (KO) emerges as the discipline within Information Science aimed at studying and developing systems for organizing units of knowledge (concepts) and for information retrieval when represented by these concepts.

In this discipline, studies focus on the thematic and descriptive treatment of information. The former deals with the subject matter addressed by an object, while the latter concerns the characteristics of the object. For the discipline to function, it develops tools that support knowledge representation, known as Knowledge Organization Systems (KOS), which are intended to assist information professionals in the representation of information.

Among the research efforts that contribute to the development of KOS. According to Hjørland (2017, sec. 2.1) «a domain, on the other hand, is a specialization in the division of cognitive labor that is theoretically coherent or socially institutionalized. […] domains are not ready-made divisions of the world but are dynamic, developing, and theory dependent».

One of the proposed research objectives within the group is the development of a Thesaurus on Research Data Management. To construct the thesaurus, it is essential to have in-depth knowledge of the subject area it will cover, as only after this stage can the selection of terms that will comprise the thesaurus be defined. In this context, domain analysis proves to be an important tool, as it provides an overview of how a field of knowledge or discipline develops.

As Freitas and Albuquerque (2017, 2) explain, «the theoretical-methodological approach of Domain Analysis – specifically in studies related to Information Science – has among its goals to support the research process and the construction of tools for knowledge organization». In this context, domain analysis serves to understand what is being researched and discussed regarding a particular concept.

Thus, the central question of this study is: How can the domain analysis approach, stemming from KO, be used to map the field of Data Management, gaining substantial knowledge to support the development of a thesaurus on the topic?

[1] Acronym of *Biblioteconomia, Representação, Interoperabilidade, E-science e Tecnologia*, translated Librarianship, Representation, Interoperability, E-science and Technology.

The general objective is to describe what constitutes the domain of Data Management in Brazilian Information Science. The specific objectives include investigating the scientific output in the field and identifying postgraduate-level courses and disciplines related to the topic.

Data management is a multidisciplinary field, involving various areas beyond Information Science, such as computer science, management, and applied fields like healthcare. However, the aim of this work is precisely to analyze how the topic of data management has been addressed specifically within the scope of Information Science, considering its theoretical frameworks, terminologies, and specific approaches.

We acknowledge, however, that this delimitation may result in a partial perspective, especially regarding the construction of a controlled vocabulary such as a thesaurus. In view of this, the text will emphasize the justification for this delimitation, indicating that it is a conscious methodological choice aimed at understanding the specificity of the Information Science domain within a broader field.

In this way, it is believed that research can contribute to studies on KO and information representation in data management, as well as support the development of a thesaurus on the topic.

2. Data Management

According to Borgman (2012), data is the driving force behind research in any field of knowledge. Their format will depend on their origin, purpose, and academic discipline. For example, a rock may serve as data for both History and Geology. A photograph can be considered research data in Anthropology, History, Medicine, Biodiversity, and other fields.

In this sense, research data are not generated exclusively in digital formats; notebooks, analog photographs, drawings, handwritten calculations, among others, are also considered research data, as long as they are produced within the context of a research activity.

Today, there are various ways to share data; however, data sharing alone does not ensure data reuse. Effective data management is essential. Researchers seek to ensure long-term access to their data (Shah et al. 2025).

Data management encompasses aspects related to acquisition, organization, storage, use and reuse, and preservation of data. All of this must be supported by appropriate policies and infrastructures. According to Majid, Zhang and Foo (2018, 2) «Research data management entails all activities and processes which are undertaken or done to ensure that research data is properly documented, organised, stored, archived and curated so that it is available for access, use and reuse whenever the need arises after the research has been done and reported».

It can be stated that research data management encompasses the set of activities and mechanisms that span from data capture to its reuse. Processes that ensure data organization–such as description and storage–are essential. «Data management covers all aspects related to handling, organizing, documenting, and adding value to data, and it plays a crucial role as a facilitator in data sharing processes, ensuring long-term data sustainability and accessibility» (Sayao and Sales, 2015, 5).

In turn, Kotronoulas et al. (2023, 2) offer a more simplified definition: «Crucially, the starting point is to carefully prepare a data set ready for analysis. Numbers (also known as raw data) must be put in a form that makes them suitable for analysis. This is called data management». That is, data management enables data to be used in a simplified and efficient manner.

Examples of data management activities include addressing the data lifecycle (from capture to reuse), access policies, sharing and reuse policies, data preservation actions, infrastructure selection, and data representation.

Data do not speak for themselves; therefore, KOS assist in their description within data repositories, data papers, and data journals. In this context, KO can play a key role in supporting effective data management.

3. Organization of Knowledge

Knowledge is a fundamental part of human life and is intrinsically linked to the functioning of society. Therefore, aspects such as science, religion, and politics can only be developed through shared knowledge–and for this to happen, it must be recorded. However, recorded knowledge alone does not guarantee its dissemination. It must also be organized and made accessible to all. According to Guimarães (2014, 14): «in order for this socially produced knowledge to have social use, a mediating process of organization is necessary, in which "substitutes of knowledge" are established, so that the contexts of production and use can be brought into dialogue».

In this context, KO emerges as a discipline whose aim is to ensure that knowledge is retrievable, playing a fundamental role within the fields of Information Science and Librarianship. It can be inferred that the development of these fields is grounded in the construction of information retrieval techniques, based on theories of knowledge organization. On this matter, Smiraglia (2012, 225) states: «Knowledge organization (also well-known by its acronym KO) is the domain in which the order of knowledge is both the primary paradigm for scientific investigation and the primary application in the development of systems».

Thus, based on the theories that underpin discipline, tools are created to support the representation and retrieval of information, such as KOS. For Mazzocchi (2024, sec. 1.0) «Knowledge organization system (KOS) is a ge-

neric term used for referring to a wide range of items (e.g. subject headings, thesauri, classification schemes and ontologies), which have been conceived with respect to different purposes, in distinct historical moments».

Knowledge Organization Systems have among their functions the task of assisting in thematic representation when a document needs to be described using keywords. Documents are written in natural language (the language of the document), whereas KOS provide standardized terms that can be used to represent the record consistently.

Those systems can take various forms, including thesauri, subject heading lists, glossaries, controlled vocabularies, ontologies, taxonomies, among others. Their structures vary according to their intended purpose, which may include improving information retrieval, enhancing navigation, facilitating communication among specialists, mapping knowledge domains, supporting intelligent machine processing, and more. In this way, KOS have specific characteristics that distinguish them from one another.

Therefore, KOS can support the retrieval of information within research data. For Given and Olso (2003, 160) «the principal concepts of knowledge organization and their interrelations, when construed broadly, offer a model for the intellectual work involved in preparing data for analysis».

It is not possible to understand a dataset without its exhaustive description. For the reuse of datasets made available in sharing environments to occur, a proper description is necessary for comprehension. According to Meschini and Francelin (2022, 4), «Knowledge Organization has the potential to act as a mediator between issues related to the ease of data acquisition and its complex interpretation, contributing to the semantic expansion of this data phenomenon, especially concerning scientific data».

Moreover, KOS, such as a thesaurus, can assist researchers not only during the sharing and reuse of data but also in knowledge representation when research outputs like abstracts and articles are developed.

As mentioned in the introduction of this study, the current research aims to support the BRIET research group in the development of a thesaurus on data management, with the goal of helping Brazilian researchers better represent their texts according to the topic.

To develop a thesaurus, it is necessary to understand the domain to which it will belong. Thus, the objective of this research was to carry out a domain analysis of data management within the context of Brazilian Information Science. The step-by-step process of how the research was conducted will be presented in the next section.

4. Methodological Procedures

To meet the objectives, a Domain Analysis study was conducted, developed according to Hjørland (2002), as it is an approach widely used in the field of Information Science research, specifically within the scope of Knowledge Organization. According to Smiraglia (2011, 1):

> Domain analysis is the study of the evolution of discourse around theoretical poles represented through the formal products, usually the literature, of a research community. Domain analysis is one way of generating new knowledge about the interaction of communities of scholars with information. Domain analysis of international research communities brings the promise of new comprehension of how people interact with information in different places.

For its development, several methodological steps must be undertaken. In line with the objectives proposed for the present research, the approaches suggested by Hjørland (2002) were considered. Specifically, Hjørland (2002, 422) proposes 11 approaches, which are:

> producing literature guides and subject gateways; producing special classifications and thesauri; research on indexing and retrieving specialities; empirical user studies; bibliometrical studies; historical studies; document and genre studies; epistemological and critical studies; terminological studies, LSP (languages for special purposes), discourse studies; studies of structures and institutions in scientific communication; and domain analysis in professional cognition and artificial intelligence.

For the development of the present research, two approaches were selected: bibliometric study and epistemology. The bibliometric study aims to demonstrate how scientific production occurs within a given domain and/or subject. In turn, according to Hjørland (2002), epistemological and critical studies make it possible to understand the theories, methodologies, and applications of a domain. This approach may seem more basic, but it is necessary. According to the author, when this approach is not employed, domain analysis tends to be more superficial.

For the epistemological study, academic Graduate Programs in Information Science were selected. The justification for choosing these programs is twofold: first, there are no undergraduate degrees in Information Science in Brazil; second, these programs are advanced courses that focus on the study and practice of information management, organization, retrieval, and dissemination. These programs typically address topics such as librarianship, archival science, information technologies, data analysis, and information policy. Therefore, they deal directly with information and are also focused on discussing new demands for information organization, such as the topic of data management.

To identify and select the Graduate Programs, a search was conducted on the Sucupira Platform. The Sucupira Platform is a digital platform that centralizes information from the National Graduate System. Graduate Programs fill out the platform annually with data that the government can use for planning and decision-making. The tool enables the management of this level of Brazilian education.

The search was conducted using the filters provided by the platform, which are: Major Area of Knowledge, Evaluation Area, and Field of Knowledge. For this research, the selected categories were, respectively: Major Area of Knowledge – Applied Social Sciences; Evaluation Area – Communication and Information; and Field of Knowledge – Information Science.

A total of 28 programs were retrieved, of which 3 were not yet operational, being still under development. Therefore, 25 programs were considered valid for the research. After identifying the programs, their websites were examined to determine whether they offered courses that address issues related to Data Management. The search was carried out in March 2025. From the 25 programs retrieved, 11 programs were selected that offered courses related to the theme. As part of the data analysis, the course syllabi were reviewed.

One program was excluded because it was not possible to access or read the course syllabi.

To analyze data related to postgraduate programs, thematic coding was selected. According to Medelyan (2025) coding is the process of labeling and organizing your qualitative data to identify different themes and the relationships between them. That is, thematic coding allows the researcher to structure the data in a way that facilitates its subsequent interpretation. In other words, coding is not just a technical act, but an interpretative and organizational step.

Still according to Medelyan (2025), when coding customer feedback, you assign labels to words or phrases that represent important (and recurring) themes in each response. These labels can be words, phrases, or numbers; we recommend using words or short phrases, as they're easier to remember, skim, and organize.

For this research, the first step was to read the course syllabi to gain a general understanding of the topics. Following this, the topics were divided based on the course syllabi. Then, the topics were grouped by theme, and finally, the areas covered by the courses were identified. Thus, thematic coding was performed using the course titles, following Mendelay's (2025) guidelines. The results of the final coding, with the grouping of courses and graduate programs, can be seen in the next chapter.

For the bibliometric study, the selected database was the *Base de Dados em Ciência da Informação (BRAPCI)*, a Brazilian digital infrastructure designed for the collection, preservation, and access to scientific literature in the field of Information Science. It includes a wide range of publications, such as journal

articles, conference papers, books, and book chapters, primarily from Brazilian and Latin American sources (BRAPCI n.d.). The database also includes some journals from Portugal.

Currently, the database includes over 100 journals in the field and publishes proceedings from 8 national events. It contains articles from the year 1962 to the present. The platform allows searches by fields such as author, abstract, keywords, title, and a combined search across all five fields. Filters can be applied to narrow results to Brazilian journals, foreign journals, conferences, or books and book chapters. The search also supports Boolean operators, as shown in the Figure 2.



Figure 2. Brapci home screen.

To encompass current research, a time frame from 2010 to 2024 was applied. This range allows for an updated overview of the topic. The search was conducted using the titles of the works. The choice of the title field was made because the title is considered an important element in the development of academic works. A well-structured title summarizes the main topic and the purpose of the research. It is also believed that a clearly described title improves discoverability in databases and digital libraries, as it is the first element users encounter in publications.

Two search terms were used in the database: *Data Management* and *Research Data Management*, aiming for a more precise retrieval. A total of 128 works were retrieved: 96 articles with the term *Data Management* and 32 with *Research Data Management*. After removing duplicates, 97 articles were analyzed.

For the article analysis, the title, abstract, keywords, and introduction were read. As the focus of the research is to understand the themes and approaches related to data management, other information related to bibliometric studies, such as quantitative data, authorship analysis, and collaboration networks, were not discussed. The articles were categorized by thematic areas for analysis and will be presented in the results chapter.

5. Results

The first analysis carried out was of the Graduate Programs. Of the 10 programs selected for the final stage of the research, 1 program focuses on Information Management and the remaining 9 are centered on Information Science.

Even though it is not part of the thematic coding, it is considered necessary to present a description of the research lines of the graduate programs investigated, as they offer courses that aim to contribute to their scope.

The *Information Management Program* at *UDESC (Santa Catarina State University)* focuses on the discussion of topics related to the management of organizations that deal with information, aiming to propose solutions to concrete problems. It is committed to ethics and to the economic, social, and educational role of information in citizenship, preparing professionals capable of solving complex problems in information management while generating and disseminating technical and scientific knowledge for innovation.

In turn, the *Graduate Program in Information Science* at *UFPE (Federal University of Pernambuco)* is centered on Information as its main theme, with Memory and Technology as its theoretical framework. Even so, it offers a course that covers the scope of Data Management.

The *Information Science Program* at the *Federal University of Santa Catarina (UFSC)* has "Information and Society" as its area of concentration. According to the program's official webpage, it emphasizes the study of phenomena related to core processes such as management, selection, processing, organization, retrieval, use, and dissemination of information, aiming primarily to address, explore, understand, and explain these phenomena.

The *Information Science Program* at *São Paulo State University (UNESP)* is focused on "Information, Technology, and Knowledge". It develops research that encompasses aspects of information and technology, the production and organization of information, and the management, mediation, and use of information.

Meanwhile, the *Information Science Program* at the *Federal University of Alagoas (UFAL)* focuses on "Information, Technology, and Innovation". Research in this program is directed toward the production, mediation, communication, and management of information and technological processes.

The *Information Science Program* at the *Federal University of Rio Grande do Sul (UFRGS)* concentrates on "Information, Science, and Society". Its goal is to develop research on the relationships between information and science, as well as in different social contexts, involving the discussion and analysis of contemporary problems to improve information flows and their management in various environments.

The *Information Science Program* at the *Federal University of Rio Grande do Norte (UFRN)* focuses on "Information in Contemporary Society". This concentration area aims to address trends in theories, methodologies, and practices in Information Science. It considers issues specific to post-industrial society related to the production, organization, mediation, management, dissemination, access, and use of information and knowledge.

The *Information Science Program* at the *Federal University of Ceará (UFC)* conducts research focused on the *Representation and Mediation of Information and Knowledge*. To this end, its focus is divided into studies on information, organization, communication and technology, and information, society, and culture.

The *Information Science Program* at the *Brazilian Institute for Information in Science and Technology (IBICT)* has its area of concentration in *Information and Social and Technological Mediations for Knowledge*. Therefore, researchers in the program focus on aspects of information and its social and technological mediations for knowledge.

The *Information Science Program* at the *University of São Paulo (USP)* is centered on *Information and Culture*. It deals with the relationships that characterize the processes of construction and/or reconstruction of meaning and/or cultural products when information is transformed into knowledge and the cultural product into a cultural asset. It proposes the observation of the necessary actions – within the context of cultural institutions – for information to be preserved and socially circulated (collection, selection, organization, access), and the analysis of the cultural contexts in which these processes occur and acquire social meaning.

After selecting the programs, a total of 18 courses related to the topic were identified. These courses address aspects related both to general *Data Management* and to more specific topics such as *Research Data Management*. Below is a description of what was recovered in the first reading of the course syllabi.

It is evident that the courses offered are intrinsically linked to the field of Information Science, and each program provides courses aligned with its area of concentration or scope. From an epistemological perspective of Information Science, the courses found reflect the objectives set forth in the programs' scopes, contributing to issues such as the organization, description, preservation, visualization, and use of data for information generation.

The *Information Management Program* at *Santa Catarina State University (UDESC)* offers two courses focused specifically on data management: *Data Governance and Research Data Management*. Regarding the *Data Governance* course, its objective is to understand governance concepts and their application in the context of scientific data, including Open Science and Citizen Science contexts, as well as the creation of a Scientific Data Management Plan. The *Research Data Management* course aims to discuss e-Science practices, focusing on Research Data: concepts, types, and structure. It also covers platforms for Data Management, including Research Data Repositories.

Meanwhile, the *Federal University of Pernambuco (UFPE)*, through its Graduate Program in Information Science with a focus on Memory and Technology, offers the course *Digital Curation*. The Digital Curation course provides a methodological framework for issues related to the preservation and management of data and digital information, involving the maintenance and enhancement of research data throughout its lifecycle. This concept aims to ensure the accessibility and use of research data in the future, as well as to promote its discovery, interpretation, and reuse.

The *Federal University of Santa Catarina (UFSC)* offers three courses related to the theme of data management: *Big Data and Semantic Web; Web Data Management for Information Science*; and *LGPD (General Data Protection Law) and the New Legal Framework for Innovation: Interfaces between Law and Information Science*.

The *Big Data and Semantic Web* course presents concepts and discusses aspects related to the Semantic Web and supporting technologies. It also covers the conceptual foundations of ontologies applied to information systems, practices of Linked Data, semantic information retrieval, Big Data, and associated information technologies.

The *Web Data Management for Information Science* course focuses on the characteristics of Web data, modeling of semi-structured data, access to Web data, as well as the integration, processing, analysis, and visualization of data.

The *New Legal Framework for Innovation: Interfaces between Law and Information Science* course discusses aspects related to the General Data Protection Law (LGPD), which is a law designed to establish rules for the processing of personal data, whether in physical or digital form, by individuals or legal entities.

In the Graduate Program at *São Paulo State University (UNESP)*, three courses were identified: *Information Science and the Data Lifecycle: Research Frontiers; Open Science and Research Data Management: Theory and Practice; and Data Science for Information Science*.

The course *Information Science and the Data Lifecycle: Research Frontiers* discusses the role of Information Science in the collection, processing, storage, dissemination, and preservation of data.

The course *Open Science and Research Data Management: Theory and Practice* aims to provide a theoretical deep dive into the concepts of Open Science and Research Data Management, as well as the points of convergence between these two movements. Topics discussed include the FAIR Principles, Data Librarianship, Data Management Plans, Data Stewardship, and other aspects related to research data management.

The course titled *Data Science for Information Science* covers discussions on Data Science applied to Information Science, enabling students to understand and apply the principles, methods, and techniques of Data Science within the field of Information Science. This involves studying aspects related to data collection, organization, analysis, and interpretation, allowing the extraction of relevant knowledge that contributes to efficient information management.

In the Graduate Program in Information Science at the *Federal University of Alagoas (UFAL)*, two courses were identified: *Open Science and Digital Scientific Marketing and Open Data and E-Governance*.

The Open Science and Digital Scientific Marketing course discusses collaborative and open scientific initiatives and practices, such as e-science and research data. It also presents concepts of digital marketing applied to scientific products.

The *Open Data and E-Governance* course covers concepts related to open government data and electronic governance.

In the Graduate Program in Information Science at the *Federal University of Rio Grande do Sul (UFRGS)*, three courses were identified: *Open Science and Scientific Reproducibility; Scientific Communication; and Scientific Data Management*.

The course Open Science and Scientific Reproducibility studies aspects of Open Science and data openness, data publication, FAIR (Findable, Accessible, Interoperable, Reusable) Principles, and other aspects of research data management.

The *Scientific Communication* course addresses issues related to open science and open access, as well as topics concerning open research data.

The *Scientific Data Management* course aims to investigate the conceptual and practical relationship between the research process and data handling across all knowledge areas. It includes the study of the main international organizations involved in developing policies and tools for scientific data management, types of scientific data, and the development of data management plans as well as preservation and reuse processes in repositories. The course also covers the scientific data lifecycle in relation to Open Science and the mastery of fundamental aspects necessary for the digital preservation of scientific data, in addition to studying the role of information professionals.

The Graduate Program in Information Science at the *University of São Paulo (USP)* offers a course on the topic titled *Data Visualization and Analytical Cul-*

ture. This course focuses on the introductory exploration of software and tools for creating dashboards based on usage scenarios, as well as theoretical reflection on these informational objects within an analytical culture that emerges from the mediation of Information and Communication Technologies.

In the Graduate Program in Information Science at the *Federal University of Rio Grande do Norte (UFRN)*, the course *Scientific Communication and Open Science* was identified. This course discusses Open Science, presenting concepts, schools of thought, principles, and collaborative and open scientific practices. It also covers open data, research data management, and Open Science policies.

The *Federal University of Ceará (UFC)*, in its Graduate Program, also offers a course titled *Scientific Communication*. This course presents discussions about Open Access and Open Data.

In the Graduate Program of the *Brazilian Institute of Information for Science and Technology (IBICT)*, the course also called *Scientific Communication* addresses issues related to open data and the publication of such data.

In summary, as part of the thematic coding, the following is a description of the disciplines with the most prominent emphasis:

- Data Governance;
- Digital Curation;
- Open Science;
- Data Visualization;
- Personal Data Protection (LGPD);
- Scientific Communication;
- Repositories and Preservation;
- Semantic Web and Big Data;
- Data Lifecycle.

The initial codes are based on the first reading and thematic grouping. These are:

Excerpt	**Code**
"understand Governance concepts and their application in the context of scientific data"	Data governance
"Open Science and Citizen Science"	Open Science
"Creation of the Scientific Data Management Plan"	Data Management Plan (DMP)
"Digital curation" / "maintenance and valorization of research data throughout the lifecycle"	Digital Curation / Data Lifecycle
"Big Data and Semantic Web"	Big Data / Web Semântica
"FAIR principles"	FAIR principles

"Data Stewardship, Data Librarianship"	Data librarianship
"Data preservation and reuse in repositories"	Data reuse
"General Data Protection Law"	Personal data protection (LGPD)
"Data Visualization and Dashboards"	Data visualization

Table 1. Initial codes, extracted directly from the data.

After the analysis carried out in the coding, it is possible to perceive the central themes and their groupings, together with the postgraduate programs.

Universidade	**Disciplinas**	**Main Themes (Thematic Coding)**
UDESC (Santa Catarina State University)	- Data Governance - Research Data Management	- Data Governance - Open Science - Citizen Science - Data Management Plan - Data Platforms and Repositories
UFPE (Federal University of Pernambuco)	- Curadoria Digital	- Digital curation - Data preservation and lifecycle - Data accessibility and reuse
UFSC (Federal University of Santa Catarina)	- Big Data and Semantic Web - Web Data Management for Information Science - LGPD and the New Legal Framework for Innovation	- Big Data - Semantic Web - Ontologies - Web Data and Analytics - General Data Protection Law (LGPD)
UNESP (São Paulo State University)	- Information Science and the Data Lifecycle - Open Science and Research Data Management - Data Science for Information Science	- Data Lifecycle - Open Science - FAIR Principles - Data Librarianship and Data Stewardship - Data Science - Data-Based Information Management

Universidade	Disciplinas	Main Themes (Thematic Coding)
UFAL (Federal University of Alagoas)	- Open Science and Digital Scientific Marketing - Open Data and E-Governance	- Open Science - Digital Science Marketing - Open Government Data - E-Governance
UFRGS (Federal University of Rio Grande do Sul)	- Open Science and Scientific Reproducibility - Scientific Communication - Scientific Data Management	- Open Science - Scientific Communication - FAIR Principles - Repositories and Preservation - Information Professionals in Data Management
USP (University of São Paulo)	- Data Visualization and Analytical Culture	- Data visualization - Analytical culture - Tools and dashboards
UFRN (Federal University of Rio Grande do Norte)	- Scientific Communication and Open Science	- Open Science - Open Data - Open Science Policies
UFC (Federal University of Ceará)	- Scientific Communication	- Scientific communication - Open access - Open data
IBICT (Brazilian Institute of Information for Science and Technology)	- Scientific Communication	- Scientific communication - Publication and access to open data

Table 2. Subjects and Topics Related to Data Management in Postgraduate Programs in Information Science – Brazil (2025).

The disciplines demonstrate how the field of Information Science is integrating:

- Emerging epistemologies (open science, FAIR data);
- Digital infrastructures (repositories, platforms, big data);
- Legal and social demands (LGPD, reproducibility, innovation);
- Interdisciplinary practices (curation, data science, law).

This integration reveals an evolution of the field toward a systemic management of scientific information, strengthening its social, scientific, and political relevance.

It is noticeable that, when analyzing the courses offered by these programs, such perspectives promote discussion on epistemological innovations in Information Science. It is possible to observe the evolution of the field alongside the advancement in knowledge availability, as seen in data management. Moreover, the field has been utilizing its theories and studies involving issues of organization, processing, archiving, preservation, and use to support theoretical questions about this new domain, which is Data Management.

The second approach developed in the present research was the analysis of bibliographic production, which was conducted using one of the specific databases for Brazilian Information Science, *Brapci*.

When analyzing the scientific production of the domain, it becomes evident that it reflects what is being studied in the Graduate Programs. The topics retrieved in this part of the research present study results strongly linked to the themes of the courses, such as Open Science, research data, data visualization, the General Data Protection Law, FAIR Principles, digital curation, as well as the role of the information professional.

It can be observed that the main topics developed are closely connected to research data management, including Open Science, e-Science, the data lifecycle, data repositories, FAIR Principles, and data management plans. These aspects are recurrent in both the scholarly works analyzed and the courses offered by the Graduate Programs.

Within the theme of data curation, several studies address digital preservation, especially within data repositories, and refer to the data lifecycle. Still within the scope of curation, works were also found that deal with the use of metadata for data curation and for the description of data in repositories.

Regarding the theme of the FAIR Principles, these are directly linked to research focused on data repositories, as they serve as guiding principles for such technologies. Additionally, there are studies on FAIR objects and FAIR data management plans.

Additional studies were also found that focus on the *data librarian* and the role of the librarian in data management–an aspect that was not identified in the analysis of the Graduate Programs. Since *Librarianship and Information Science* are closely linked, it is believed that the presence of bibliographic production on this topic reflects what is being studied in undergraduate Librarianship programs in Brazil.

After reading the articles, one can find aspects that reflect the epistemology of Information Science. Such as the idea that research data are not merely technical records, but fundamental elements for the construction, validation, and circulation of scientific knowledge. There is a premise that knowledge needs to be organized, described, and accessible to be legitimized.

The bibliographic production also includes articles discussing *data visualization, services related to data management, data quality, Big Data*, and the

General Data Protection Law (LGPD). This Brazilian law, approved in 2018, regulates privacy and the use/processing of personal data in Brazil.

Based on the results presented, the following Figure 2 attempts to synthesize the findings by combining epistemological study and the bibliometric study.

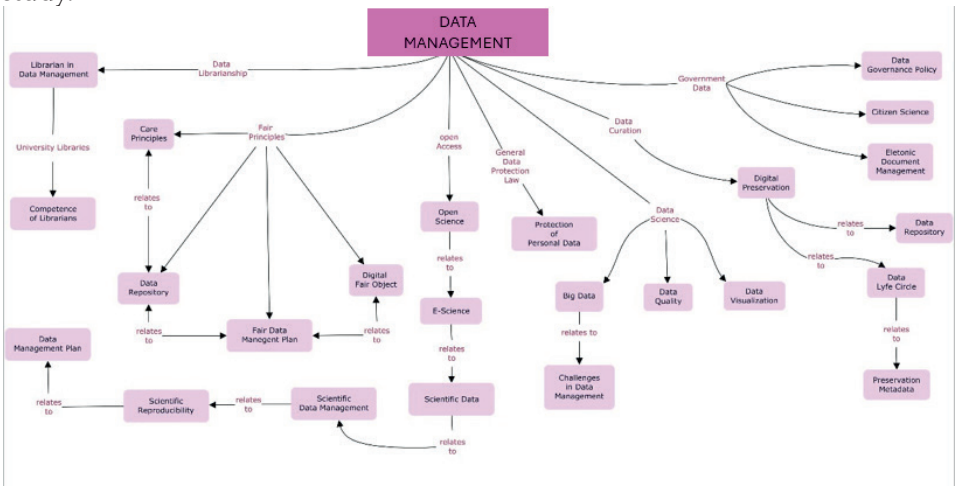


Figure 2. Conceptual map about Data Management.

The choice to develop a *concept map* stem from its ability to provide an initial visualization of the field, which can contribute to the creation of the *Briet Group's thesaurus*, since a thesaurus is composed of terms organized in a hierarchical structure (Grupo Briet n.d).

When observing the map, the central field is *Data Management*, which branches into six subfields. These, in turn, are formed by secondary topics that constitute the research landscape. Thus, it can be inferred that the field of Data Management in Brazilian Information Science is composed of seven main themes: *Data Librarianship, FAIR Principles, Open Access, Data Science, General Data Protection Law, Data Curation*, and *Government Data*. These are supported by secondary, yet significant, topics that enrich discussions both within Graduate Programs and in the construction of the bibliographic field.

This demonstrates that Brazilian Information Science has been evolving in its *epistemological aspects*, embracing new nuances in the creation, organization, and dissemination of knowledge.

6. Final Considerations

Domain analysis proves to be an important tool for the understanding and development of fields such as *Data Management*. It is hoped that this work

will effectively contribute to the progress proposed by the *Briet research group*. Based on the analysis of the results, it is evident that, in the investigated context, *Brazilian Information Science* is primarily focused on *Research Data Management*.

It was observed that both the courses and the articles analyzed have strong *epistemological connections* with Information Science, as it is a discipline that deals with various aspects of information. This demonstrates that Information Science is in tune with the new demands of *knowledge production and organization*.

It is acknowledged that choosing the field of Information Science as the corpus to support the research and, subsequently, to serve as a basis for the construction of a thesaurus entails certain methodological and conceptual challenges. These implications can be observed in several aspects, such as:

- *Incomplete domain representation*: Considering that the field of Data Management has a multidisciplinary nature, there is a risk of excluding relevant terms, which would compromise the conceptual scope of the thesaurus.
- *Interoperability issues*: Thesauri and controlled vocabularies are often used in information systems that require interoperability. The absence of standardization or alignment between systems may hinder communication among them.
- *Low usefulness for users:* An effective thesaurus must comprehensively reflect users' information retrieval needs. Otherwise, its practical applicability may be significantly reduced.
- *Weaknesses in evolution and updating:* There is a risk that the thesaurus may become obsolete if it remains restricted to a specific thematic approach, making it difficult to adapt to new informational demands.

To mitigate these risks, it is recommended that, before finalizing the thesaurus, the team responsible for its development carefully consider these implications. It is essential to adopt strategies that enable the inclusion of the various subfields that make up the Data Management domain, thereby enriching the thesaurus's base corpus and reducing the occurrence of conceptual gaps.

Additionally, it is suggested that the thesaurus be subject to periodic reviews and updates to ensure its relevance, adaptability, and alignment with the evolution of the thematic domain.

References

Alonso-Arévalo, Julio. 2019. "La gestión de datos de investigación en el horizonte de las bibliotecas universitarias y de investigación." *Cuadernos de Documentación Multimedia* 30: 75-88. https://doi.org/10.5209/CDMU.62806.

Borgman, Christine L. 2012. "The conundrum of sharing research data." *Journal of the American Society for Information Science and Technology* 63, no. 6 (June):1059-78. https://doi.org/10.1002/asi.22634.

BRAPCI (Base de Dados em Ciência da Informação). n.d. "Sobre a Brapci!" Last Accessed November 2, 2024. https://brapci.inf.br/about/brapci.

Freitas, Lidiane Marques, and Ana Cristina de Albuquerque. 2017. "As abordagens da análise de domínio como aporte metodológico para a classificação arquivística." *Tendencias da Pesquisa Brasileira em Ciência da Informação* 10(2): 1-19. https://revistas.ancib.org/tpbci/article/view/424.

Given, Lisa M., and Hope A. Olson. 2003. "Knowledge organization in research: A conceptual model for organizing data." *Library & Information Science Research* 25, no. 2 (Summer): 157-76. https://doi.org/10.1016/S0740-8188(03)00005-7.

Grupo Briet. n.d. "Briet: Biblioteconomia, Representação, Interoperabilidade, E-science e Tecnologia." Last Accessed December 16, 2024. https://grupobriet.com/.

Guimarães, José Augusto Chaves. 2014. "Análise de domínio como perspectiva metodológica em organização da informação." *Ciência da Informação* 43, no. 1 (jan.-abr.):13-21. https://doi.org/10.18225/ci.inf.v43i1.1415.

Hjørland, Birger. 2002. "Domain analysis in information science: eleven approaches traditional as well as innovative." *Journal of Documentation* 58, no. 4 (August):422-62. http://dx.doi.org/10.1108/00220410210431136.

Hjørland, Birger. 2017. "Domain analysis." *Knowledge Organization* 44 (6):436-64. https://doi.org/10.5771/0943-7444-2017-6-436. Also available in ISKO Encyclopedia of Knowledge Organization, edited by Birger Hjørland and Claudio Gnoli, https://www.isko.org/cyclo/domain_analysis.

Kotronoulas, Grigorios, Suzana Miguel, Maura Dowling, et al. 2023. "An Overview of the Fundamentals of Data Management, Analysis, and Interpretation in Quantitative Research." *Seminars in oncology Nursing* 39, no. 2 (April):e161398. https://doi.org/10.1016/j.soncn.2023.151398.

Majid, Shaheen, Shubert Foo, and Xue Zhag. 2018. "Research Data Management by Academics and Researchers: Perceptions, Knowledge and Practices". In *Maturity and Innovation in Digital Libraries*. *ICADL 2018*, edited by Milena Dobreva, Annika Hinze, and Maja Žumer. *Lectures Notes in Computer Science* 11279. Cham: Springer. https://link.springer.com/chapter/10.1007/978-3-030-04257-8_16.

Mazzocchi, Fulvio. 2024. "Knowledge Organization System (KOS)." *ISKO Encyclopedia of Knowledge Organization*, November 11. https://www.isko.org/cyclo/kos.

Medelyan, Alyona. 2025. "Coding Qualitative Data: How to Guide." Last updated February 24, 2025. https://getthematic.com/insights/coding-qualitative-data.

Meschini, Fabio Orsi, and Marivalde Moacir Francelin. 2022. "Organização do Conhecimento e suas contribuições em um contexto Big Data". *Transinformação*, Campinas, 34:e210075. https://doi.org/10.1590/2318-0889202234e210075.

Plataforma Sucupira. n.d. "Informações e serviços." Last Accessed March 1, 2025. https://sucupira.capes.gov.br/.

Sayão, Luis Fernando, and Luana Farias Sales. 2015. *Guia de Gestão de Dados de Pesquisa para Bibliotecários e Pesquisadores*. CNEN/IEN. https://www.aben.com.br/Arquivos/420/420.pdf.

Shah, Naimat Ullah, Salman Bin Naeem and Robina Bhatti. 2025. "Digital data sets management in university libraries: challenges and opportunities". *Global Knowledge, Memory and Communication* 74(1-2):446-62. https://doi.org/10.1108/GKMC-06-2022-0150.

Smiraglia, Richard P. 2011. "Domain coherence within Knowledge Organization: People, Interacting Theoretically, Across Geopolitical and Cultural Boundaries." In *Proceedings of the 39th Annual Conference of CAIS*, edited by Pam Mckenzie, Catherine Johnson, and Sarah Stevenson. University of New Brunswick. https://doi.org/10.29173/cais601.

Smiraglia, Richard P. 2012. "Knowledge organization: some trends in an emergent domain." *El profesional de la información* 21, no. 3 (mayo-junio):225-27. https://doi.org/10.3145/epi.2012.may.01.

AIDAinformazioni

Rivista semestrale di Scienze dell'Informazione

Anno 43

N. 3-4 – luglio-dicembre 2025

Contributi

Giovanna Aracri, Antonietta Folino, Maria Teresa Guaglianone, Marco Manna, Identification of thesaurus relationships with the support of ChatGPT and Gemini

Alessandra Bonsignori, Terminologia giuridica e organizzazione della conoscenza

Maria Teresa Chiaravalloti, Maria Teresa Guaglianone, Elena Cardillo, Metadatazione delle tipologie di documenti clinici: un approccio alla standardizzazione mediante LOINC

Carla Beatriz Marques Felipe, Carla Maria Martellote Viola, Raimunda Fernanda dos Santos, Domain Analysis on Data Management

Claudio Gnoli, Emanuela Valenzano, Che faccetta sei?

Manuela Grillo, Strumenti di convergenza tra archivi e biblioteche: i thesauri

Bianca La Manna, Rappresentare gli atti notarili medievali in Linked Open Data (LOD)

Claudia Lanza, L'univocità e la variabilità terminologica a supporto della rappresentatività delle risorse semantiche di organizzazione e rappresentazione della conoscenza specialistica

Larissa Moraes Martins, Carlos Eduardo da Silva Carvalho, Ana Cristina de Albuquerque, Mario Barité, Possibilities for using cultural warrant to complement literary work

Bianca Therezinha Carvalho Panisset, Lucia Maria Velloso de Oliveira, Leticia Souza da Costa Sampaio, Camilla Campoi de Sobral, Identification of types of documents and knowledge organization: transparency and access in records management

SoYoung Yoon, Claudio Gnoli, Ziyoung Park, Applying Integrative Levels Classification to Korean KOS Registry

Emanuela Nicole Donato, Grazia Serratore, Maria Taverniti, Maria Teresa Chiaravalloti, L'evoluzione della struttura documentale nelle cartelle cliniche psichiatriche dell'archivio storico del manicomio di Girifalco (1881-1978)

Lorenza Talarico, Manuscripts in the Digital Age

Maria Lígia Triques, Ananda Fernanda de Jesus, Mario Barité, José Eduardo Santarem Segundo, Ana Cristina de Albuquerque, Vocabularies in the digital context

Marcin Trzmielewski, Context-Driven Use of Automatic Tools in Medical Ontology Design: A Case Study in Allergology

Carla Maria Martellote Viola, Carla Beatriz Marques Felipe, Exploring the relationships between the Cross-Cutting Women's Agenda - Multi-Year Plan 2024-2027 and the Knowledge Organization of legislative proposals under processing on women in Brazil

Elena Cencetti, Camillo Carlo Pellizzari di San Girolamo, Elisabetta Viti, Modellizzazione delle conoscenze: Thesaurus Nuovo soggettario e Wikidata, due strumenti a confronto

Pontish Yeramyan, Alexander Murzaku, Nick Uva, Curt Anderson, The transferability of brilliance

Ana Flávia Dias Zammataro, Ana Cristina de Albuquerque, Giulia Crippa, Lucia Sardo, Carlo Bianchini, Maria Lígia Triques, ISKO Italy in perspective

mundaneum

In copertina
Disegno di Paul Otlet, Collections Mundaneum, centre d'Archives, Mons (Belgique).

ISBN 979-12-5965-639-1

ISSN 1121-0095