Journal of Theoretical and Applied Information Technology
30th November 2024. Vol.102. No. 22

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

INSIGHTS INTO METADATA COMPONENTS: A SYSTEMATIC REVIEW OF ENTERPRISE DATA CATALOGS

ERIC WIJAYA, TANTY OKTAVIA

Information Systems Management Department, BINUS Graduate Program – Master of Information Systems, Bina Nusantara University

Jakarta, Indonesia 11480

E-mail: [1]eric.wijaya001@binus.ac.id, [2]toktavia@binus.edu

ABSTRACT

In the contemporary data-centric business environment, proficient metadata and data documentation processes are essential for organizations aiming to maximize the value of their data assets. This data-driven architecture, when coupled with Business Intelligence (BI) tools, promotes data democratization, allowing stakeholders throughout the business to utilize it. This research examines the exploration of metadata elements on the development of BI systems. A thorough literature research and a preliminary analysis are done to comprehend the landscape of metadata classification. This study delves into existing research on metadata and data catalogue management in enterprises, utilizing a systematic literature review (SLR) to identify specific metadata components. The SLR results not only describe the functions of each metadata component, but also provide practical guidance on how to adopt them, making them helpful insights for firms wishing to enhance their data management. Along with the bibliometric study investigates trends and partnerships among metadata, providing further information on efficient metadata implementation. These findings have implications for firms looking to improve their data processes and achieve a competitive advantage by providing new insights into management tactics, opening the way for future research in metadata and data catalogue systems.

Keywords: *Metadata, Systematic Literature Review, Bibliometric, Catalogue System*

1. INTRODUCTION

The primary function of big data is to facilitate managerial decision-making [1]. Valuable insights that are essential for decision-making can be obtained through big data analytics, which can disclose previously unknown findings in big data discovery endeavors [1]. Building a comprehensive data warehouse is insufficient for an organization to completely leverage its data. They must guarantee that the data is accessible to all data consumers in each department, thereby facilitating the development of data-informed decisions. This data consumer is not limited to conventional data professionals (e.g., business intelligence specialists, data managers, data analysts, or data scientists), but also includes employees who utilize self-service business intelligence tools in their daily work, who are occasionally referred to as "data citizens" [2]. The FAIR principles, which entail that data should be Findable, Accessible, Interoperable, and Reusable, are used to conceptualize the challenges associated with extracting data from a variety of sources and preparing it for use [3].

One of the prerequisites for ensuring that data is FAIR is data documentation, which involves the inventory and detailed description of data assets to aid data workers, particularly those who are not data specialists, in locating and comprehending the data [2]. For this requirement, metadata, which is data about data [4], [5], is an obvious choice for data documentation [9]. The data can be documented in a variety of ways using metadata, including the meaning of its content, information on data quality or security, and data lifecycle aspects [6].

Metadata serves as the fundamental component of data catalogs due to its comprehensive form [7]. A systematic literature analysis of metadata and data catalog implementation in enterprise settings is employed in this research to gain a more comprehensive understanding of the function of each metadata component. Moreover, this research will broaden the current metadata classification to address a wider range of metadata implementation aspects, in accordance with the previous studies in metadata classification [7], [8]. This study seeks to enhance the scientific corpus of knowledge in the metadata components by analyzing the comprehensive aspects of metadata.

2. THEORITICAL BACKGROUND

2.1 Metadata

Metadata is the most fundamental component of the FAIR data principles. In order to distinguish between data and metadata, the FAIR data governing principles of the original paper defined data as all digital resources, including software tools and not just data in its conventional sense. In contrast, metadata is any description of a resource that is intended to facilitate the FAIR principles of that resource [3]. A prevalent occurrence is the misinterpretation of unstandardized or unfamiliar metadata as regular data, a phenomenon referred to as "one person's metadata is another person's data." By treating the data/metadata pair in isolation, FAIR principles address this confusion by asserting that metadata is the descriptor and data is the object being described within the context of that pair [9]. Ensure that data is FAIR (Findable, Accessible, Interoperable, and Reusable) by conducting an exhaustive data documentation process that involves the systematic inventory and detailed description of data assets. This procedure is designed to aid data workers, particularly those who are not data experts, in the locating and understanding of the data [2]. Consequently, metadata, which is frequently referred to as "data about data" [4], [5], is an indispensable instrument for satisfying this prerequisite [2]. Metadata functions as a method for documenting a variety of aspects of the data, including its content's significance, security measures, and even aspects of its lifecycle [6].

By investigating the overall administration of metadata, metadata classification can be further developed. The utilization of standardized ontologies [11] [12] in the development of the metadata model is one of the metadata management components [2]. The interoperability of the metadata and the adherence to the FAIR principles are guaranteed by standardized metadata formats [7], [13], [14]. Optimization of metadata implementation and maintenance necessitates automation [14]–[16]. The automation process would significantly enhance the quality of metadata and its monitoring, as emphasized in [14] and [7]. Continuous metadata enrichment can be accomplished in a collaborative manner when combined with manual reviews and revisions [7], [7]. By offering a UI application or Business Intelligent system that enables users to review, update, or delete metadata as required, these seamless integration and implementation can be achieved [8].

2.2 FAIR

The FAIR data principles are designed to offer a framework for enhancing the reusability, accessibility, interoperability, and findability of digital resources [9]. The principles, which were introduced in 2016, have been widely cited, endorsed, and adopted by a variety of parties. The FAIRification procedure, which is the workflow for transforming data into FAIR, was initially devised for a series of "Bring Your Own Data" (BYOD) workshops [10]. The initial objective was to develop a methodology that is domain-independent and can be used in a diverse array of FAIRification initiatives. One of the generic workflows proposed by [18] was divided into three phases:

1. Pre-FAIRification, which involves the identification of the FAIRification objective, as well as the analysis of data and metadata.
2. FAIRification, which entails the definition of a semantic data and metadata model, the linking of data and metadata, and the hosting of FAIR data.
3. Post-FAIRification, which involves the evaluation of the FAIR data

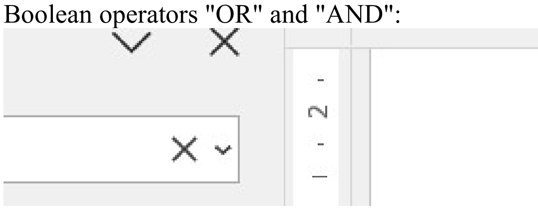
3. METHODOLOGY

This study employs a Systematic Literature Review (SLR) methodology, which is distinguished by its comprehensive and structured approach to the review and consolidation of research evidence within a specific field. The method in question guarantees transparency, reproducibility, and bias reduction, as emphasized by [17]. The validity of a study is contingent upon the proper selection of a database, which ensures adequate coverage of the area under investigation [18], [19]. Therefore, the selection of appropriate data sources is essential for the collection of the most prospective literature pertinent to the study. We have taken into account Scopus (www.scopus.com), the largest multidisciplinary database with over 40,000 reviewed journals, as well as four other sources for data extraction: Web

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

of Science (www.webofscience.com), EI Compendex (www.engineeringvillage.com), IEEE Xplore (ieeexplore.ieee.org), and ACM (dl.acm.org), in accordance with the recommendation from [20].

Relevant keywords and their respective alternatives are incorporated into the search query, as this research concentrates on the implementation of metadata in enterprise environments. The following is the resultant combination of the Boolean operators "OR" and "AND":



The initial analysis is bibliometrics analysis, which was derived from the systematic literature review conducted by [21]. Bibliometric analysis is advantageous when conducting a thorough examination of a substantial quantity of scientific data while simultaneously generating a substantial research impact [22]. Scientific research that is not specific to an enterprise setting, but includes metadata and data catalog keywords, is analyzed to offer a comprehensive perspective on metadata research. We then selected a research paper that was specifically tailored to the enterprise context. The only papers that are accessible and have a metadata component as the primary topic are subsequently reviewed manually.

4. RESULT

This sub-section gives the results of the bibliometric analyses for the 455 papers, accompanied by the commentary. The data's graphical depiction is produced using VOS Viewer, facilitating network relation research.

This analysis concentrates on publications from 2009, as the subject of data cataloging and data management began to acquire prominence in 2016 [23]. Table 2 illustrates the annual distribution of published articles. The table indicates that articles on data catalog and metadata are dispersed throughout the year.

Citation analysis is employed to identify the seminal publications within a specific scientific domain. This approach evaluates the influence of a publication by examining its citation count, facilitating the identification of the most important works within a study domain [22]. The bibliometric analysis of metadata and data management reveals important insights into intellectual processes and academic effect through the examination of citation counts in this sector. Table 1 presents a compilation of significant works in metadata and data cataloguing, ranked by the frequency of citations each publication has garnered. The cutoff criterion was established at 40 citations, yielding 13 articles from 2009 to 2021. The analysis reveals that significant research in metadata management is closely linked to data management. Of the 13 papers, six pertain to data or metadata management, and five concern data lake management. This research illustrates the essential role of metadata in the development of a data lake or data warehouse.

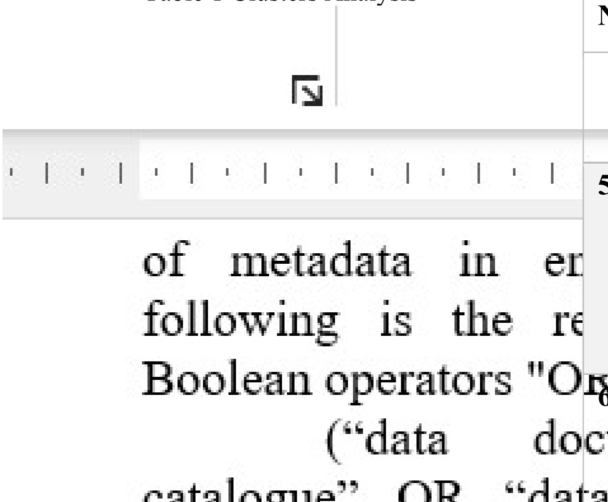
Keyword co-occurrence analysis is a type of co-word analysis that functions as a proxy for analyzing the actual content of the publication [22]. This analysis also allows us to identify the relationships between topics in a specific research field by examining the written content of the publication [24]. The VOSViewer application is employed to assist in the generation of the network, as it has the ability to generate and color code comparable topics from the list of publications. To capture the intended topics from the author's perspective, we also examine the author keyword in place of the index keyword for this analysis.

From the 455 publications, the authors referenced a total of 1,182 keywords. In order to enhance the visualization of only the pertinent keyword, a threshold limit is implemented. The number of keywords that are generated decreases as the exclusion threshold increases, as illustrated in Table 2. The threshold was established at a minimum of three occurrences for this analysis. The co-occurrence network of 92 author keywords.

According to the author keyword co-occurrence network result, the network is composed of 9 main clusters, each of which contains a minimum of 5 items, and 5 minor clusters, each of which contains one or three items. Each of these clusters encompasses two primary subjects in the field of metadata research: metadata management and metadata itself. Here is a list of the 9 major clusters and their primary items:

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1 Clusters Analysis


of metadata in er
following is the re
Boolean operators "Or
("data doc
catalogue" OR "data

In the subsequent section, the qualitative analysis results of metadata classification are detailed, along with the components of each group. Using 16 accessible publications on the implementation of metadata and data catalogs in enterprise settings, a qualitative systematic literature review is conducted. Information regarding metadata components is extracted and categorized through the review of all 16 papers. Each paper that was examined during this procedure is detailed in Table 2.

Table 2 Studies Reviewed

No	Author (Year)	Year	Title
1	Ehrlinger L.; Schrott J.; Melichar M.; Kirchmayr N.; Wöß W.	2021	Data Catalogs: A Systematic Literature Review and Guidelines to Implementation
2	Qi X.	2021	Research on Enterprise Data Governance Based on Knowledge Map
3	Yu Q. (2010)	2010	Metadata integration architecture in enterprise data warehouse system
4	Schutz C.; Schrefl M.	2014	Customization of Domain-Specific Reference Models for Data Warehouses
5	Fujita Y.; Naono K.; Hanai T.	2012	Proposal and evaluation of metadata management method for eDiscovery
6	Yan Y.; McLane T.	2012	Metadata management and revision history tracking for spatial data and GIS map figures
7	Dela Cruz N.; Schiefelbein P.; Anderson K.; Hallock P.; Barden D	2010	ORM and MDM/MMS: Integration in an enterprise level conceptual data model
8	van Helvoirt S.; Weigand H.	2015	Operationalizing data governance via multi-level metadata management
9	Seng J.-L.; Wong Z.	2012	An intelligent XML-based multidimensional data cube exchange
10	Vnuk L.; Koronios A.; Gao J. (2011)	2011	Enterprise metadata management: Conceptions, issues and capabilities
11	Labadie C.; Legner C.; Eurich M.; Fadler M.	2020	FAIR Enough? Enhancing the Usage of Enterprise Data with Data Catalogs
12	Labadie C.; Eurich M.; Legner C.	2020	Empowering data consumers to work with

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

No	Author (Year)	Year	Title
			data: Data documentation for the enterprise context
13	Chelmis C.; Zhao J.; Sorathia V.; Agarwal S.; Prasanna V.	2013	Toward an automatic metadata management framework for smart oil fields
14	Shanmugam S.; Seshadri G.	2016	Aspects of Data Cataloguing for Enterprise Data Platforms
15	Dzyubanenko A.A.; Rabin A.V.	2022	Hybrid client-server implementation and microservice architecture of automatic documentation analysis software
16	Petrik D.; Untermann A.; Baars H. (2024)	2024	Functional Requirements for Enterprise Data Catalogs: A Systematic Literature Review

We employ metadata categories as proposed by [8] for the classification of each component, a method that is frequently employed in other research studies on the classification of metadata components. Those metadata categories are administrative metadata, descriptive metadata, and structural metadata. The metadata management category is added as an umbrella category to incorporate the metadata management criteria identified in several studies, as per the bibliometric analysis' findings. Consequently, there are 17 metadata components categorized into four metadata categories, which will be elaborated upon in the subsequent sections.

4.1 Descriptive metadata

The descriptive metadata offers users a comprehensive overview of the data, including its title, concise description, and potential utilization statistics. This information is essential for locating the data that is most suitable for their analysis and requirements. The following components are included in this category:

1. The most fundamental form of metadata is the title and description. This information contains definitions and descriptions of the content of a data [13]. It can be tagged or placed at four levels: dataset level, record level, entity level, and column level, dependent on the context of the utilization [25]. A descriptive title and a sufficient description are necessary for a data knowledge to be more consumable, whether by manual searching by data consumers or by other systems through indexing [26].

2. The anticipated data value and the degree of adherence to the standardized format are inextricably linked due to the data formats [25]. Format standardization is a critical component of evaluating the quality of data in terms of semantics or validity, particularly in the context of multidimensional data exchange [27]. When it comes to thousand separators and fractional portions, data format is frequently observed in numerical data, phone numbers, currencies, and dates [18], [21].

3. Another factor in the maintenance and evaluation of enterprise data quality is data ranges [14]. As per [18], data ranges enable data users to identify any deviations within a specific dataset, as well as its anticipated limits, as they are depicted in the data. Data range attributes are inherent to numerical and periodical data, which is why they are readily quantifiable. On the other hand, a custom rule will be required for string format data. For instance, the length of postal code data is limited to five or six digits.

4. Usage statistics, as their name implies, monitor the consumption of specific data [2]. It is a component of the operational metadata, which includes run-time information relevant to the use or query of the data [29]. The usage statistics metadata is classified as descriptive metadata due to its derivative nature, which can only be envisaged after the data being monitored has been established [13].

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4.2 Administrative metadata

Technical information, such as accessibility, sensitivity, and data lineage information, are captured by administrative metadata. As specified by the research that elaborates on these aspects in their study, each component can be described as follows:

1. The knowledge of data transformation and mutation series is carried by data lineage as it passes through numerous applications or the construction of data cubes and data marts [25]. This information is essential for the purpose of tracing the steps taken by each sequence in order to guarantee the quality, reliability, and authenticity of the data [25]. Data lineage information is frequently represented in lineage graphs to illustrate the data migration across multiple systems [30].

2. Data governance is inextricably linked to data sensitivity and accessibility. PII, or personally identifiable information, should be designated as sensitive information, and access to such data may be restricted [7]. Compliance with regulations and the prevention of data misconduct necessitate the regulation of sensitive information and its access [26]. Access controls and access policies must be established in multi-level metadata management to accommodate intricate data access rules in business environments [29].

3. Another component of data governance is the administrative designation of responsible individuals or teams to the data. The administrative aspects of this component included the assignment of duties and responsibilities, as well as the establishment of procedures [14]. In order to ensure that data sets are managed and used responsibly, it is necessary to establish clear data custody [26].

4. The auditing capability that is essential in the data security management aspect is facilitated by data structure or attribute modification tracing [13]. The name of the creator/modifier and the time of the operation are the fundamental details that should be prioritized when monitoring the modifications to extant data and the creation of new data [31]. Furthermore, the audit record may incorporate revision or modification notes to facilitate tracing [32].

4.3 Structural metadata

Through the provision of data linkage, data origin, and the business context encircling a data, structural metadata assists users in comprehending the relationship between the data. Each component in this category is described in detail as follows:

1. The interconnectivity among data objects and the metadata itself is supported by data linkage and relationship. Due to the interdependence of enterprise business processes, the data from one process may be associated with one or more data from other processes [25]. This relationship is as prevalent as either a key-based relationship, a master-instance relationship, or a source-contributor relationship [25]. The relationship between conceptual, physical, and logical data views should also be documented by an enterprise metadata model [13], [14], [26].

2. Data origin is the earliest level at which the provenance information of a data can be traced back to its original system source [7], [26]. Additionally, data origin metadata may be implemented to conserve information regarding the hardware and software locations, which is advantageous for users to identify the precise origin of the data [12].

3. The business glossary or data business context assists in the alignment and mapping of the system-oriented nature of a data with its applicable business domain [25]. This information is essential for a comprehensive comprehension of specific business terms [26], [30]. The target audience for metadata or data catalogue implementation is not limited to IT specialists; it also includes business users who are involved in data utilization for the purpose of improving it [7].

4. A prevalent practice in data warehouse implementation is the capture of technical data changes during the Extract, Transform, Load (ETL) procedures through data transformation and calculation [29]. Derivation or customization rules are frequently included in this information, which is used to generate the data in a specific data mart [2] or during the calculation of business key performance indicators (KPIs) [2].

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4.4 Metadata management

Metadata management is the process of generating and maintaining metadata, which includes the preservation of metadata information to ensure its veracity, as well as the utilization of the built-in metadata catalogue by users. This category serves as an umbrella that incorporates the comprehensive management process of metadata implementation. Its components include:

1. Standardized ontologies or metadata standard usage guarantee consistency in the management of metadata implementations through standardization [32]. The key component of data interoperability and integration within metadata systems is the consistent implementation of metadata standards [13], [33]. XML, RDF, and JSON-LD are frequently employed as metadata formats due to their machine-readable nature [33]. Four metadata standards that are pertinent to the enterprise context and data catalogs were retained, as determined by the research conducted by [2] in the development of a custom metadata model. The Dublin Core Schema, the Data Catalog Vocabulary, the ISO 11179-3 Metadata Registry Metamodel and Basic Attributes, and the Common Warehouse Metamodel are the four standards in question.

2. The optimization of the data catalog and metadata implementation, as well as their ongoing maintenance, is facilitated by the implementation of automation in metadata management [14]. Automating the management of metadata information and maintenance can be used to maintain the complexity and completeness of detailed documentation, thereby reducing the risk of failure and the need for manual user involvement [14]. Additionally, an intelligent automation system can assist in the reduction of time-consuming and manual tasks associated with data discovery and analysis [26].

3. The automation process can be complemented by the establishment of a metadata revision and review process. Users are anticipated to be able to review, revise, or delete metadata as necessary, contingent upon their role [15]. It is a critical component of the automation process, as it is essential for the ongoing enhancement of metadata administration and data catalog [7]. The collaborative character of the management process will allow users to collaborate in the enrichment of metadata [14], reduce silos [2], [30], and foster synergies between various user groups [26].

4. The key to a data's findability is robust search functionality, which enables data consumers to determine the existence of the data and its location [14]. Keywords, fragmentary metadata identifiers, or full text are among the fundamental entries that users must include when conducting data discovery [26]. This capability will be further enhanced by integrating the search functionality that is specifically designed for business context, which will enable the identification of pertinent data objects based on the business terms [2].

5. Metadata accessibility refers to the availability of the data catalog and metadata to consumers [14]. In order to facilitate digital interactions for both technical and business users and to enable for the visual presentation of metadata information, digital user interfaces are frequently developed [26]. This functionality can be further enhanced by offering user interface customization, multilingual support [13], and application programming interfaces (APIs) that can be accessed by other systems [26], [28].

5. CONCLUSIONS

The SLR implemented in this investigation meticulously extracted and synthesized metadata components from the existing literature that pertains to metadata and data catalogues in enterprise contexts. This study offers a nuanced comprehension of the complex operations of metadata systems within organizational frameworks by defining the roles and providing implementation guidance for each identified component. This comprehensive analysis can provide organizations with valuable insights into the optimization of their metadata and data catalogue implementations, thereby enhancing data utilization and gaining a competitive advantage in their respective industries.

Additionally, the bibliometric study that was conducted as part of this research illuminated the collaboration patterns, sub-topics, and emerging trends within the metadata research domain. This analysis not only enhances our comprehension of the current state of metadata research, but also provides valuable insights for future research directions and collaborative opportunities among academicians and practitioners in this field.

The results of this study have significant practical implications for organizations that are attempting to enhance the implementation of their metadata and data catalogues. This research provides organizations with the knowledge and tools required to improve their data management practices, thereby enabling them to more effectively leverage their data assets to achieve strategic business objectives, by elucidating guidelines and strategies for metadata management in enterprise settings.

This study enhances the current metadata classification framework by incorporating an additional segment dedicated to management strategies and provides a comprehensive overview of metadata and data catalogue components, thereby contributing to the existing body of knowledge on a theoretical level. This innovative segment is essential for the encapsulation of the comprehensive approach necessary for the successful implementation of metadata and data catalogues, thereby furthering theoretical understanding and providing a foundation for future research endeavours in this field.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the financial support of DRPTM DIKTI Kemendikbudristek under the scheme of Hibah Penelitian Tesis Magister with main contract number: 0459/E5/PG.02.00/2024, on 30 Mei 2024, Contract Number of LLDIKTI 3-Binus: 105/E5/PG.02.00.PL/2024; 784/LL3/AL.04/2024; 092/VRRTT/VI/2024, on 8th August 2024

REFERENCES

[1] S. Jeble, S. Kumari, and Y. Patil, "Role of big data in decision making," *Oper. Supply Chain Manag.*, 2018, doi: 10.31387/oscm0300198.

[2] C. Labadie, M. Eurich, and C. Legner, "Empowering data consumers to work with data: Data documentation for the enterprise context," 2020. doi: 10.30844/wi_2020_m7.

[3] M. Wilkinson, M. Dumontier, and I. Aalbersberg, "The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 3 (2016)," *Sci. data*, 2016.

[4] R. Roszkiewicz, "Enterprise metadata management: How consolidation simplifies control," *J. Digit. Asset Manag.*, 2010, doi: 10.1057/dam.2010.32.

[5] W. H. Inmon, B. K. O'Neil, and L. Fryman, *Business Metadata: Capturing Enterprise Knowledge*. 2007. doi: 10.1016/B978-0-12-373726-7.X5001-9.

[6] R. Eichler, C. Giebler, C. Gröger, H. Schwarz, and B. Mitschang, "Handle - a generic metadata model for data lakes," 2020. doi: 10.1007/978-3-030-59065-9_7.

[7] L. Ehrlinger, J. Schrott, M. Melichar, N. Kirchmayr, and W. Wöß, "Data Catalogs: A Systematic Literature Review and Guidelines to Implementation," 2021. doi: 10.1007/978-3-030-87101-7_15.

[8] J. (NISO) Riley, "Understanding metadata: what is metadata, and what is it for?: a primer," *NISO Prim. Ser.*, 2017.

[9] A. Jacobsen *et al.*, "Fair principles: Interpretations and implementation considerations," *Data Intelligence*. 2020. doi: 10.1162/dint_r_00024.

[10] C. Carta *et al.*, "The FAIRification of data and the potential of FAIR resources demonstrated in practice at the Rome Bring Your Own Data workshop," 2017.

[11] A. Jacobsen *et al.*, "A generic workflow for the data fairification process," *Data Intell.*, 2020, doi: 10.1162/dint_a_00028.

[12] R. Gabriel, T. Hoppe, and A. Pastwa, "Classification of metadata categories in data warehousing - A generic approach," 2010.

[13] L. Vnuk, A. Koronios, and J. Gao, "Enterprise metadata management: Conceptions, issues and capabilities," 2011.

[14] C. Labadie, C. Legner, M. Eurich, and M. Fadler, "FAIR Enough? Enhancing the Usage of Enterprise Data with Data Catalogs," 2020. doi: 10.1109/CBI49978.2020.00029.

[15] C. Chelmis, J. Zhao, V. Sorathia, S. Agarwal, and V. Prasanna, "Toward an automatic metadata management framework for smart oil fields," 2013. doi: 10.2118/153271-PA.

[16] C. Schutz and M. Schrefl, "Customization of Domain-Specific Reference Models for Data Warehouses," 2014. doi: 10.1109/EDOC.2014.18.

[17] K. Petersen, S. Vakkalanka, and L. Kuzniarz, "Guidelines for conducting systematic mapping studies in software engineering: An update," *Inf. Softw. Technol.*, 2015, doi: 10.1016/j.infsof.2015.03.007.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[18] E. S. Vieira and J. A. N. F. Gomes, “A comparison of Scopus and Web of science for a typical university,” *Scientometrics*, 2009, doi: 10.1007/s11192-009-2178-0.

[19] L. Chen, M. A. Babar, and H. Zhang, “Towards an Evidence-Based Understanding of Electronic Data Sources,” 2010. doi: 10.14236/ewic/ease2010.17.

[20] A. Carrera-Rivera, W. Ochoa, F. Larrinaga, and G. Lasa, “How-to conduct a systematic literature review: A quick guide for computer science research,” *MethodsX*, 2022, doi: 10.1016/j.mex.2022.101895.

[21] F. A. Rachman and T. Oktavia, “A Systematic Literature Review of Critical Success Factors for User Experience in Gamification of Work,” *J. Syst. Manag. Sci.*, 2024, doi: 10.33168/JSMS.2024.0214.

[22] N. Donthu, S. Kumar, D. Mukherjee, N. Pandey, and W. M. Lim, “How to conduct a bibliometric analysis: An overview and guidelines,” *J. Bus. Res.*, 2021, doi: 10.1016/j.jbusres.2021.04.070.

[23] E. Zaidi, G. De Simoni, R. Edjlali, and A. D. Duncan, “Data Catalogs Are the New Black in Data Management and Analytics,” *Gartner*, 2017.

[24] K. J. Emich, S. Kumar, L. Lu, K. Norder, and N. Pandey, “Mapping 50 Years of Small Group Research Through Small Group Research,” *Small Gr. Res.*, 2020, doi: 10.1177/1046496420934541.

[25] S. Shanmugam and G. Seshadri, “Aspects of Data Cataloguing for Enterprise Data Platforms,” 2016. doi: 10.1109/BigDataSecurity-HPSC-IDS.2016.52.

[26] D. Petrik, A. Untermann, and H. Baars, “Functional Requirements for Enterprise Data Catalogs: A Systematic Literature Review,” 2024. doi: 10.1007/978-3-031-53227-6_1.

[27] J. L. Seng and Z. Wong, “An intelligent XML-based multidimensional data cube exchange,” *Expert Syst. Appl.*, 2012, doi: 10.1016/j.eswa.2012.01.069.

[28] A. A. Dzyubanenko and A. V. Rabin, “Hybrid client-server implementation and microservice architecture of automatic

[29] S. van Helvoirt and H. Weigand, “Operationalizing data governance via multi-level metadata management,” 2015. doi: 10.1007/978-3-319-25013-7_13.

[30] N. Dela Cruz, P. Schiefelbein, K. Anderson, P. Hallock, and D. Barden, “ORM and MDM/MMS: Integration in an enterprise level conceptual data model,” 2010. doi: 10.1007/978-3-642-16961-8_67.

[31] Y. Fujita, K. Naono, and T. Hanai, “Proposal and evaluation of metadata management method for eDiscovery,” 2012. doi: 10.1109/SRII.2012.88.

[32] Y. Yan and T. McLane, “Metadata management and revision history tracking for spatial data and GIS map figures,” 2012. doi: 10.1145/2345316.2345357.

[33] E. Quimbert, K. Jeffery, C. Martens, P. Martin, and Z. Zhao, “Data cataloguing,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2020. doi: 10.1007/978-3-030-52829-4_8.