ISSN: 2723-9535

Available online at www.HighTechJournal.org

HighTech and Innovation Journal

Vol. 6, No. 1, March, 2025

Strategic Metadata Implementation: A Catalyst for Enhanced BI Systems and Organizational Effectiveness

Tanty Oktavia [1*], Eric Wijaya [1]

[1] *Information Systems Management Department, BINUS Graduate Program, Master of Information Systems Management, Bina Nusantara University, Jakarta 11480, Indonesia.*

Received 31 October 2024; Revised 25 January 2025; Accepted 07 February 2025; Published 01 March 2025

Abstract

In today's data-driven business landscape, robust metadata and data documentation practices are essential for enterprises aiming to maximize their data assets. When integrated with Business Intelligence (BI) systems, this architecture empowers data democratization, allowing widespread utilization by stakeholders across the organization. This research explores the critical role of metadata in shaping Business Intelligence (BI) systems and organizational effectiveness within today's data-driven business landscape. Through a systematic literature review, a preliminary study, a quantitative survey with 318 responses, and a focus group discussion, the study identifies key metadata components influencing BI systems effectiveness and organizational outcomes. Findings indicate a direct and positive impact of BI systems effectiveness on organizational effectiveness. Certain metadata components exhibit direct positive effects on both BI systems and organizational effectiveness. The research underscores the importance of strategic metadata implementation for enterprises seeking to optimize data-driven decision-making processes. Overall, the study provides practical implications for organizations and contributes valuable insights to the understanding of metadata's role in enhancing enterprise effectiveness.

Keywords: Metadata; Data Documentation; Business Intelligence; Data Analytics; Organizational Effectiveness.

1. Introduction

The advent of the big data era has empowered organizations to accumulate vast datasets, leveraging data analytics and advanced data science for strategic insights. As highlighted by Peyton [1], the data landscape within organizations undergoes a doubling every five years, leading to an abundance of redundant and inconsistent data. Effectively managing this data deluge often necessitates the construction of an enterprise data warehouse, serving as a foundational requirement to aggregate and centralize all available data. This consolidation is pivotal in providing data consumers within the organization with a unified environment for discovering and processing diverse datasets.

However, the mere existence of a comprehensive data warehouse does not guarantee optimal utilization. To harness the full potential of their data, organizations must ensure accessibility for every data consumer across various departments. These consumers encompass not only traditional data experts, including business intelligence specialists, data managers, analysts, and scientists, but also extend to employees utilizing self-service business intelligence tools—often referred to as "data citizens" [2]. Overcoming challenges associated with extracting and refining data from diverse sources, ensuring they align with the FAIR principles (Findable, Accessible, Interoperable, Reusable), becomes pivotal in facilitating a cohesive and effective data utilization strategy [3].

* Corresponding author: toktavia@binus.edu

http://dx.doi.org/10.28991/HIJ-2025-06-01-02

One crucial requirement for achieving FAIR data is the meticulous documentation of data assets, necessitating comprehensive inventories and detailed descriptions tailored to assist data workers, including those who may lack expertise in the field [2]. In addressing this need, metadata emerges as a natural and instrumental candidate, defined as data about data [4, 5]. As highlighted by Labadie et al. [2], metadata plays a pivotal role in providing essential data documentation. Its capacity as data about data proves invaluable in documenting various facets of the data, including elucidating the meaning of its content, providing insights into data quality or security parameters, and detailing aspects of the data lifecycle [6].

Metadata itself comprises a vast number of components, ranging from simple table titles and column descriptions [7] to more complex technical representations of a data lineage [8]. Conducting a holistic study that captures all metadata components would necessitate a classification effort. Numerous previous studies [7–11] have tried to classify metadata into several categories. Some authors differentiate metadata into functional and technical metadata, while others try to differentiate based on the purpose and the systematic nature of the metadata [10]. For example, Riley [7] classified metadata into 3 general categories: descriptive metadata, administrative metadata, and structural metadata.

The positive effect of metadata implementation in enterprise settings will be evaluated in the context of business intelligence (BI) systems and organizational effectiveness. BI systems assume a central role in leveraging constructed metadata to support business analytics and decision-making processes [12, 13]. The imperative for metadata management, essential in rendering data FAIR and ensuring data quality, resonates across diverse industries. The study by Ehrenmann et al. [14] underscores the absence of a one-size-fits-all solution optimal for different industries. Notably, BI applications equipped with high-quality metadata emerge as a critical factor, ensuring user understanding and fostering trust in organizational data [13]. Consequently, the efficiency and productivity of a company stand to improve when the data provided through BI is perceived as valuable [15, 16].

Well-documented data, complete with the BI systems to provide easy access for users, is the key to achieving FAIR data. High quality and completeness of metadata is the key for effective BI systems [13, 16]. In general, metadata can be categorized into three categories, mainly descriptive metadata, administrative metadata, and structural metadata [7, 9, 11]. The fourth category in this research is obtained by extending the previous categories and incorporating important aspects of metadata management found in numerous studies. By focusing on the metadata factors and their relationship to BI systems and organizational effectiveness, this study attempts to address the following research questions:

- What is the relationship between metadata factors and BI systems effectiveness?
- What is the relationship between BI systems and organizational effectiveness?
- Does a BI system mediate the relationship between metadata factors and organizational effectiveness?

This research combines quantitative and qualitative approaches. A ‘post-positivist’ quantitative methodology is employed to test thirteen hypotheses concerning metadata, BI systems, and organizational effectiveness. Conversely, an ‘interpretivist’ qualitative research methodology is utilized to capture diverse perspectives in interpreting the results [17]. The structure of the paper is outlined as follows. In the next section, the research contribution is presented. In section 3, we presented the systematic literature review (SLR) process to define the categorization of metadata components. Section 4 explains the preliminary study conducted to identify the gap between current metadata implementation and the expectations of the data users. Section 5 presents the methodology used in this research study. The results of the quantitative empirical study will be presented in Section 6, while the qualitative study to supplement the findings will be provided in Section 7. A discussion of these findings will be presented in Section 8. Section 9 provides the limitations and future direction. Lastly, Section 10 concludes the paper.

2. Research Contribution

The findings of this research carry significant practical implications for both organizations seeking to enhance their business intelligence systems and overall effectiveness. The first finding is consistent with previous studies [16, 18] in which the focus on improving BI systems effectiveness will also improve organizational effectiveness. The insights gained from the study also suggest actionable strategies for optimizing the utilization of metadata components. For instance, the strategic implementation of metadata management is critical in improving the BI systems and organizational effectiveness. This could be achieved by providing a user interface application that facilitates data searching to improve the findability aspect of data [19]. Implementing the automation and collaborative review processes in maintaining the metadata will also ensure the scalability and continuous improvement of a metadata management process [9, 20–22]. Other metadata components, shown to also impact BI systems or organizations’ effectiveness, are crucial in improving data-driven decision-making in the organization.

On a theoretical level, this study contributes to the existing body of knowledge on the metadata components [8–11] by segmenting each component based on its characteristics and usability. This categorization allows the broad definition of metadata to be reviewed separately. Combined with the refinement of existing theory adopted from previous studies

[13, 16, 18], the relationship between each metadata component group, BI systems, and organizational effectiveness can be understood more deeply. Moreover, the preliminary study of this research prompts further exploration into the discrepancy between the current metadata implementation strategy in an enterprise setting and the usability expectation from the data users. As found by Foshay et al. [13], while data users might already have a positive perception of the metadata's usefulness, the current metadata provided is still far from satisfactory. To further enrich the findings, a qualitative study was conducted through a focus group discussion. This research thus lays the foundation for continued academic discourse and exploration in the dynamic field of metadata.

3. Systematic Literature Review (SLR)

Several previous studies have been conducted to group metadata components. To better understand the groupings of the metadata components, a systematic literature review was carried out. The following sections discuss the SLR steps in detail.

3.1. Literature Database Selection

The selection of appropriate data sources is critical for collecting the most potential literature relevant to the study, and the validity of a study depends on the proper selection of a database, ensuring adequate coverage of the area under investigation [23, 24]. Following the recommendation from [25], we have considered Scopus (www.scopus.com), which is the largest multidisciplinary database with more than 40,000 reviewed journals, along with four other sources for data extraction: Web of Science (www.webofscience.com), El Compendex (www.engineeringvillage.com), IEEE Xplore (ieeexplore.ieee.org), and ACM (dl.acm.org).

3.2. Search Strings

Since this research focuses on metadata implementation in enterprise environments, relevant keywords are added to the search string as well as their respective alternatives. Below is the resulting combination with the Boolean “OR” and “AND”:

(“data documentation” OR “data catalogue” OR “data catalog”) OR (“metadata management” AND “data” AND “metadata”) AND (“commerce” OR “ecommerce” OR “e-commerce” OR “retail” OR “industry” OR “enterprise”).

3.3. Inclusion and Exclusion Criteria

The inclusion and exclusion criteria are designed to refine the selection of literature for review, ensuring that studies pertinent to the research topic, which focuses on metadata component implementation in an enterprise setting, are incorporated while maintaining relevance. To maintain the practical relevance of data catalog applications and capture the evolution of research focus, papers published before 2009 were excluded, as the concepts of data catalog and data management began gaining significant attention in 2016 [26]. Below are the inclusion criteria:

(1) Document type is limited to conference papers and articles.

(2) Selected literature should be published in English.

(3) Studies published under the computer science and business management subject area.

The exclusion criteria are as follows:

(1) Studies whose abstract is not related to the enterprise setting.

(2) Research with inaccessible documents.

(3) Studies that have not mentioned any of the metadata aspects.

Figure 1 shows the resulting paper documents after applying each of the inclusion and exclusion criteria. The results of 15 studies were reviewed to obtain the metadata components classification in enterprise settings.

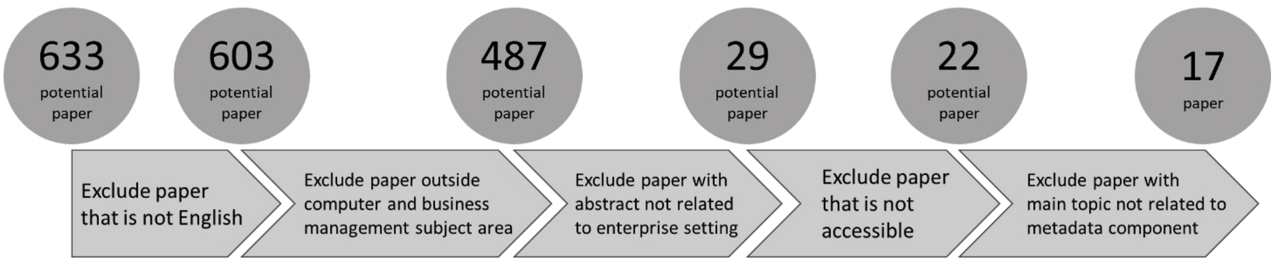

Figure 1. Number of Paper Resulted of each Criteria

3.4. Review Result

The details of each paper reviewed are presented in Table 1. It can be seen from the graph that publications on the topics of metadata in enterprise settings that align with our study criteria are distributed across the year, with a maximum of 3 publications from 2010 to 2012 and in 2021.

Table 1. Studies Reviewed

No.	Author	Title	Year
1	Ehrlinger et al. [9]	Data Catalogs: A Systematic Literature Review and Guidelines to Implementation	2021
2	Qi [27]	Research on Enterprise Data Governance Based on Knowledge Map	2021
3	Yu [28]	Metadata integration architecture in enterprise data warehouse system	2010
4	Schutz & Schrefl [29]	Customization of Domain-Specific Reference Models for Data Warehouses	2014
5	Fujita et al. [30]	Proposal and evaluation of metadata management method for eDiscovery	2012
6	Yan & McLane [31]	Metadata management and revision history tracking for spatial data and GIS map figures	2012
7	Dela Cruz et al. [21]	ORM and MDM/MMS: Integration in an enterprise level conceptual data model	2010
8	van Helvoirt & Weigand [32]	Operationalizing data governance via multi-level metadata management	2015
9	Seng & Wong [33]	An intelligent XML-based multidimensional data cube exchange	2012
10	Vnuk et al. [34]	Enterprise metadata management: Conceptions, issues and capabilities	2011
11	Labadie et al. [19]	FAIR Enough? Enhancing the Usage of Enterprise Data with Data Catalogs	2020
12	Labadie et al. [2]	Empowering data consumers to work with data: Data documentation for the enterprise context	2020
13	Chelmis et al. [20]	Toward an automatic metadata management framework for smart oil fields	2013
14	Shanmugam & Seshadri [8]	Aspects of Data Cataloguing for Enterprise Data Platforms	2016
15	Petrik et al. [35]	Functional Requirements for Enterprise Data Catalogs: A Systematic Literature Review	2024
16	Nizamis et al. [36]	Introducing an Enhanced Metadata Broker for Manufacturing Data Spaces	2023
17	Eichler et al. [37]	Introducing the enterprise data marketplace: a platform for democratizing company data	2023

To enhance the categorization and extraction of various metadata components from the literature, we employed three established metadata categories utilized in prior studies [7, 9, 11]: descriptive metadata, administrative metadata, and structural metadata. Additionally, recognizing the significance of standardized metadata languages, as emphasized in multiple studies [2, 9, 29, 33], and acknowledging the discussions on the automation of metadata management [9, 20–22], we introduced a metadata management category group to encompass aspects related to the effective administration of metadata implementation. Table 2 presents the metadata components grouping for this study alongside other studies in metadata groupings, while comprehensive details of each metadata component, including their respective mentions in each paper, are provided in Table 3.

Table 2. Metadata Components Grouping

Metadata Components Group	Quimbert et al. [11]	Shanmugam & Seshadri [8]	Gabriel et al. [10]
Descriptive	Descriptive Metadata	Data format or data ranges Data reliability	Terminology Data quality
Administrative	Administrative Metadata	Data lineage Technical context Data sensitivity and accessibility	Metadata history
Structural	Structural metadata	Data system relationship Data linkage and relationships Business context	Organization reference Data structure and data meaning System reference Data transformation
Management			Data analysis

Table 3. Metadata Components from Literature Review

Metadata Components	Paper																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Descriptive: title and description	Y			Y		Y	Y	Y	Y	Y	Y	Y				Y	Y
Descriptive: data formats	Y							Y	Y					Y		Y	Y
Descriptive: data ranges	Y							Y		Y	Y			Y	Y	Y	Y
Descriptive: Usage statistics	Y							Y		Y	Y	Y				Y	Y
Administrative: Data lineage	Y		Y	Y		Y	Y			Y	Y	Y		Y	Y		Y
Administrative: Data sensitivity and accessibility	Y	Y				Y		Y		Y	Y	Y		Y	Y		Y
Administrative: Data structure or attribute modification tracing				Y	Y	Y		Y	Y	Y		Y					
Administrative: Assignment of responsible person/team to the data	Y							Y			Y	Y		Y	Y	Y	Y
Structure: Data linkage and relationship	Y		Y	Y		Y	Y			Y	Y	Y		Y	Y		
Structure: Data origin	Y			Y					Y	Y	Y	Y		Y	Y	Y	
Structure: Data business context	Y		Y	Y			Y	Y		Y	Y	Y		Y	Y		Y
Structure: Data transformation and calculation	Y			Y			Y	Y	Y	Y	Y	Y					
Manage: Create with standardized ontologies	Y			Y		Y			Y	Y	Y	Y			Y	Y	
Manage: Automation in metadata management	Y			Y		Y	Y				Y		Y		Y		
Manage: Search Data from business context	Y		Y	Y			Y	Y		Y	Y	Y	Y		Y	Y	Y
Manage: Metadata revision and review process	Y		Y		Y	Y			Y	Y	Y	Y	Y		Y		Y
Manage: Metadata accessibility			Y								Y	Y	Y		Y		

Based on the results of the SLR, we have identified four overarching categories of metadata that will serve as the foundation for the preliminary analysis and empirical study in this research. Three of these categories were adopted from prior studies, with each individual component extracted from the reviewed literature. Additionally, one supplementary category is introduced to encompass metadata management criteria identified in several studies. The final four categories are:

(1) Descriptive metadata, provides general information about the data, which includes the title, and brief description, and might extend to usage statistics which will help users find the appropriate data for their analysis purpose and needs.

(2) Administrative metadata, captures the technical information of a data, which includes the accessibility, sensitivity, as well as the data lineage information of a data.

(3) Structural metadata, helps users to understand the relationship between the data by providing the data linkage, data origin, as well as the business context surrounding the data.

(4) Metadata management, involves the process of creating and managing metadata, including maintenance to ensure the accuracy of metadata information, as well as how users can utilize the built metadata catalog.

4. Preliminary Study

To better understand the need for metadata implementation in an enterprise setting, a preliminary study was conducted. The preliminary study was designed based on the metadata components from the systematic literature findings and was deployed to identify any gap between the current implementation in enterprise settings and the expectations of the data users.

4.1. Study Items

Each individual metadata component from the SLR results was provided as the preliminary study items. Table 4 shows the details of each item, and during the implementation, a visual aid for items in descriptive metadata, administrative metadata, and structural metadata is provided to clarify the components being reviewed. A visual aid for metadata management components is not provided since the items are mostly contextual.

Table 4. Preliminary Study Items

Expectation	Implementation
Descriptive Metadata (DM) – Table Description	
The availability of descriptions on a table can assist me in finding the needed data.	The data tables in my workplace are well-described.
DM – Column Description	
The availability of descriptions for a column in a table can help me in finding the needed data.	The columns of data tables in my workplace are well-described.
DM – Data Formats	
The availability of information on the data type in a table column can assist me in finding the needed data.	The implementation of metadata in the company where I work has effectively provided information on the data type in table columns.
DM – Data Ranges	
The availability of information on the distribution of empty data in a table column can help me assess the data quality of that column.	The implementation of metadata in the company where I work has effectively provided a depiction of the distribution of empty data.
DM – Data Preview	
The availability of data previews in a table column can help me find the needed data.	The implementation of metadata in the company where I work has effectively provided data previews.
DM – Column Usage Statistics	
The availability of usage statistics for a table can help me find the needed data.	The implementation of metadata in the company where I work has effectively provided usage statistics for tables.
DM – Table Usage Statistics	
The availability of usage statistics for a column can help me find the needed data.	The implementation of metadata in the company where I work has effectively provided usage statistics for columns.
Administrative Metadata (AM) – Table Data Lineage	
The availability of data lineage for a table helps me in finding the needed data.	The implementation of metadata in the company where I work has effectively provided data lineage for tables.
AM – Column Data Lineage	
The availability of data lineage for a column in a table helps me find the needed data.	The implementation of metadata in the company where I work has effectively provided data lineage for a column in tables.
AM – Data Sensitivity	
The availability of information about sensitive data descriptions helps me in using sensitive data.	The implementation of metadata in the company where I work has effectively provided descriptions of sensitive data.
AM – Data Accessibility	
The availability of accessibility information for data ensures the relevance of data that I can use.	The implementation of metadata in the company where I work has effectively provided accessibility information.
AM – Modification Tracking	
The availability of changes in metadata information helps me ensure the quality of available metadata.	The implementation of metadata in the company where I work effectively displays changes in metadata information.
AM – Data Ownership	
The assignment of personnel or a team responsible for a group of information ensures the quality of metadata is maintained.	The implementation of metadata in the company where I work already includes the assignment of personnel or a team responsible for a group of information.
Structural Metadata (SM) – Data Transformation	
The availability of technical details regarding the calculation/aggregation process of a data column helps me find the needed data.	The implementation of metadata in the company where I work has effectively provided information on the calculation/aggregation process.
SM – Data Relationship	
The availability of information on the relationships and connections between tables helps me find the needed data.	The implementation of metadata in the company where I work has effectively provided information on the relationships and connections between tables.
SM – Table Data Origin	
The availability of source data information for a table helps me in finding the needed data.	The implementation of metadata in the company where I work has effectively provided source data information for a table.
SM – Column Data Origin	
The availability of source data information for a column helps me in finding the needed data.	The implementation of metadata in the company where I work has effectively provided source data information for a column.
SM – Application Data Origin	
The availability of source data information for an application helps me in finding the needed data.	The implementation of metadata in the company where I work has effectively provided source data information for an application.

Expectation	Implementation
SM – Business Context	
The availability of business context information helps me in finding the needed data.	The implementation of metadata in the company where I work has effectively provided business context information.
Metadata management (MM) – Standardized Ontologies	
The creation of a metadata catalog using standard language and structure ensures the quality of metadata is maintained.	The creation of a metadata catalog in the company where I work already utilizes standard language and structure.
MM – Metadata Automation	
The automation process in creating a metadata catalog ensures the quality of metadata is maintained.	The creation of a metadata catalog in the company where I work has already implemented an automation process.
MM – Revision and Review	
Collaboration in revising and reviewing metadata ensures the quality of metadata is maintained.	The metadata catalog management in the company where I work has already implemented a collaborative process for revising and reviewing metadata.
MM – Metadata UI Accessibility	
The availability of an application with a user interface provides easy access to the built metadata catalog.	The metadata catalog built in the company where I work already has an application with a user interface.
MM – Data Findability	
The availability of search functions based on name, description, and business context can help me find the needed data.	The implementation of metadata features in the company where I work has effectively provided search functions based on name, description, and business context.

4.2. Data Collection Technique

The preliminary study was conducted through an online quantitative survey for one week asking about the current implementation and the expectations of each metadata component in the four metadata groups. The survey also used a Likert scale of 1 (strongly disagree) to 6 (strongly agree) for the statement of the implementation effectiveness as well as the usability expectation of each component. A six-point Likert scale was used as it provides more comparable reliability and construct validity than a five-point Likert scale [38].

4.3. Population and Sampling

The preliminary study was conducted for one week and served to identify any implementation gap for each of the metadata components in enterprise settings. For this reason, the sample size for this preliminary study was considerably small. A total of 33 respondents were collected at the end of the period, and a t-test was conducted to identify any significant gap in the implementation.

4.4. Results and Discussion

Figure 2 shows the overall implementation and expectation of each component group, including the expectation gap. From the result, it can be concluded that the metadata component implementation in the enterprise does not meet the expectations of the respondents participating in the preliminary study. This could be seen from the 1-point expectation gap for every metadata group as well as the p-value from the t-test showing the difference is significant. Structural metadata, which holds data relationships and business context, has the lowest implementation score and the highest expectation gap.

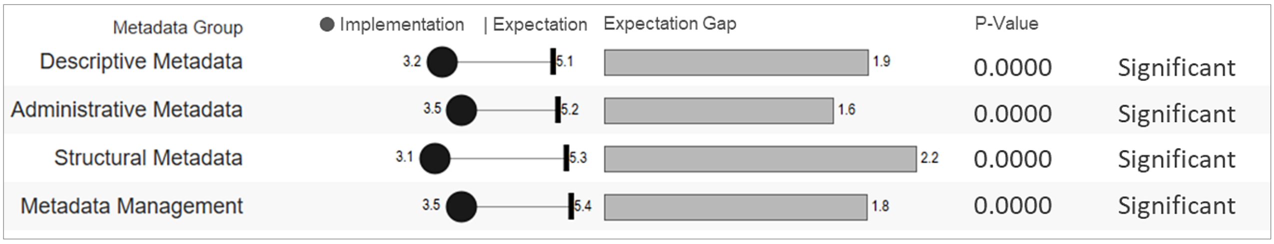

Figure 2. Implementation and Expectation Gap of Each Component Group

The detailed result for each metadata component can be seen in Figure 3. The t-test revealed significant differences between the implementation and expectation for all metadata components. By looking at each individual item, a component with the biggest expected improvement as well as its importance level can be identified. In the descriptive metadata group, this component would be the data range information with the lowest implementation point. For the administrative metadata group, column data lineage, which holds the information on the data flow of certain columns, has the highest implementation expectation gap and the lowest implementation point. The component with the highest expectation gap in the structural metadata group, which is also the highest among all components, is application data origin. It is quite understandable since the capability to identify the data source of a specific UI element will greatly improve the data findability aspects.

Metadata Group	Metadata Component	Implementation	Expectation	Expectation Gap	P-Value	
Descriptive Metadata	Table Description	3.5	5.5	2.1	0.0000	Significant
	Column Description	3.4	5.5	2.0	0.0000	Significant
	Data Formats	4.3	5.1	0.8	0.0038	Significant
	Data Ranges	2.5	4.9	2.4	0.0000	Significant
	Data Preview	3.1	4.9	1.8	0.0000	Significant
	Column Usage Statistics	2.8	4.9	2.1	0.0000	Significant
	Table Usage Statistics	2.9	4.7	1.8	0.0000	Significant
Administrative Metadata	Data Ownership	3.6	5.6	2.0	0.0000	Significant
	Table Data Lineage	4.2	5.3	1.1	0.0009	Significant
	Data Sensitivity	3.4	5.1	1.7	0.0000	Significant
	Column Data Lineage	2.9	5.1	2.2	0.0000	Significant
	Modification Tracing	3.5	5.0	1.5	0.0000	Significant
	Accessibility	3.7	4.9	1.2	0.0004	Significant
Structural Metadata	Table Relationship	3.1	5.4	2.4	0.0000	Significant
	Application Data Origin	2.7	5.4		0.0000	Significant
	Table Data Origin	3.6	5.4	1.7	0.0000	Significant
	Data Transformation	3.0	5.3	2.3	0.0000	Significant
	Column Data Origin	3.3	5.2	1.9	0.0000	Significant
	Business Context	2.9	5.0	2.1	0.0000	Significant
Metadata Management	Data Findability	3.5	5.6	2.1	0.0000	Significant
	Standardized Ontologies	3.7	5.4	1.7	0.0000	Significant
	Revision and Review	3.4	5.3	1.9	0.0000	Significant
	Metada UI Accessibility	3.8	5.3	1.5	0.0000	Significant
	Metadata Automation	3.2	5.2	2.0	0.0000	Significant

Figure 3. Detail of Implementation and Expectation Gap

5. Research Methodology

5.1. Theoretical Framework

To investigate the impact of metadata quality on organizational effectiveness, this study synthesizes concepts from several research areas in information systems, including metadata quality [16] and the BI consequences model [18]. The research model of this study is presented in Figure 4.

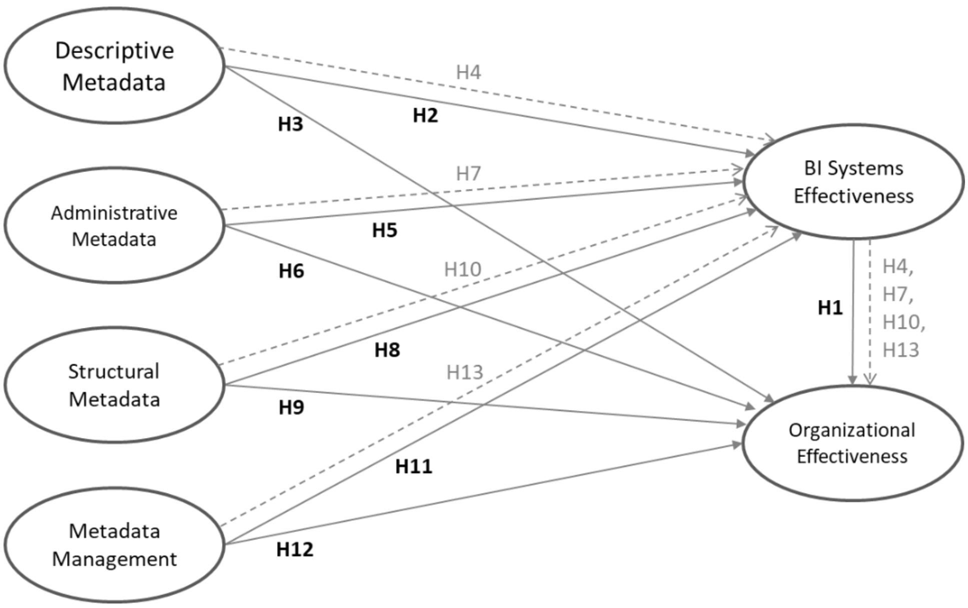

Figure 4. Research Model

5.1.1. BI Systems and Organizational Effectiveness

Several studies have shown that BI systems play roles in improving organizational effectiveness. In Masa'Deh et al. and Arefin et al. study [16, 18], the effectiveness of BI systems has been revealed to positively improve organizational effectiveness [16, 18]. BI systems provide organizations easy access to their data, in adherence to FAIR principles, and allow them to monitor, analyze, quickly adapt, and, as a result, help them in achieving objectives [16]. A study conducted by Turban et al. [39] revealed that easy access to data provided by the BI systems not only benefits the end users in the organization but also their suppliers and partners. Therefore, it is hypothesized that:

- H1. There is a positive relationship between business intelligence systems and organizational effectiveness.

5.1.2. Descriptive Metadata and BI Systems

Descriptive metadata provides general information about the data. This information includes the title and a brief description and might extend to usage statistics of data [8, 9, 11]. Consequently, descriptive metadata should help users in finding the appropriate data for their analysis purposes and needs [11]. Thus, this research hypothesized that:

- H2. Descriptive metadata will have a positive relationship with BI systems' effectiveness.
- H3. Descriptive metadata will have a positive relationship with organizational effectiveness.
- H4. BI systems' effectiveness mediates the relationship between descriptive metadata and organizational effectiveness.

5.1.3. Administrative Metadata and BI Systems

Administrative metadata captures the technical information of data. This includes data accessibility and data sensitivity, as well as the data lineage information of the data [8, 9, 11]. Data lineage acts as the supply chain of the data, which informs users of the origin of the data [16]. Because of its importance, it is hypothesized that:

- H5. Administrative metadata will have a positive relationship with BI systems' effectiveness.
- H6. Administrative metadata will have a positive relationship with organizational effectiveness.
- H7. BI systems' effectiveness mediates the relationship between administrative metadata and organizational effectiveness.

5.1.4. Structural Metadata and BI Systems

Structural metadata helps users to understand the relationship between the data by providing the data linkage, data origin, as well as the business context surrounding the data [8, 9, 11]. In general, structural metadata allows non-technical data users, such as business users, to understand the relationship between data and the business domain [2, 32]. It is hypothesized that:

- H8. Structural metadata will have a positive relationship with BI systems' effectiveness.
- H9. Structural metadata will have a positive relationship with organizational effectiveness.
- H10. BI systems' effectiveness mediates the relationship between structural metadata and organizational effectiveness.

5.1.5. Metadata Management and BI Systems

Metadata management encompasses the comprehensive implementation process of metadata. This involves the utilization of standardized ontologies, automation in metadata implementation and maintenance, and manual reviews and revisions to enhance metadata through collaboration. When combined with the availability of BI systems to facilitate user accessibility, the hypothesis is that:

- H11. Metadata management will have a positive relationship with BI systems' effectiveness.
- H12. Metadata management will have a positive relationship with organizational effectiveness.
- H13. BI systems' effectiveness mediates the relationship between metadata management and organizational effectiveness.

5.2. Variable Operationalization

Based on the theoretical framework of this study, there are 6 constructs in this research: descriptive metadata, administrative metadata, structural metadata, metadata management, BI systems effectiveness, and organizational effectiveness. Table 5 shows the items in this research along with their respective constructs.

Table 5. Construct and Measurement Items

Construct	Item	Source
Descriptive Metadata (DM)	DM1: The data description provided is complete enough to help me understand the meaning of the data. DM2: Generally, I find the data description to be accurate. DM3: Data descriptions are presented in a format that is clear and understandable.	Foshay et al. and Masa'Deh et al. [13, 16]
Administrative Metadata (AM)	AM1: Generally, the administrative information of the data provided is accurate. AM2: I get all the information I need to understand the data administration. AM3: The administrative information provided is dependable. AM4: Administrative information is presented in a format that is clear and understandable.	Foshay et al. and Masa'Deh et al. [13, 16]
Structural Metadata (SM)	SM1: The data structure information provided is accurate. SM2: I get all the information I need to understand the structure of data. SM3: The data structure information provided is dependable.	Foshay et al. and Masa'Deh et al. [13, 16]
Metadata management (MM)	MM1: I have access to a good data search facility. MM2: The BI application provided is effective in helping me to locate the data. MM3: Generally, it is easy for me to find the data I need, even if I have not used the data before.	Foshay et al. and Masa'Deh et al. [13, 16]
BI Systems Effectiveness (BI)	BI1: The BI system improves coordination with business partners/other departments. BI2: The BI system reduces the cost of transactions with business partners/other departments. BI3: The BI system improves responsiveness to/from other departments. BI4: The BI system improves the efficiency of internal processes. BI5: The BI system increases staff productivity. BI6: The BI system reduces the cost of effective decision-making. BI7: The BI system reduces operational costs. BI8: The BI system reduces customer return handling costs. BI9: The BI system reduces marketing costs. BI10: The BI system reduces time-to-market products/services.	Masa'Deh et al. and Arefin et al. [16, 18]
Organizational Effectiveness (OE)	OE1: Compared with key competitors, our company is more successful. OE2: Compared with key competitors, our company has a greater market share. OE3: Compared with key competitors, our company is growing faster. OE4: Compared with key competitors, our company is more profitable. OE5: Compared with key competitors, our company is more innovative.	Masa'Deh et al. and Arefin et al. [16, 18]

5.3. Data Collection Technique

A quantitative survey was developed using a web-based survey instrument and administered through Microsoft Forms. The choice of an online survey was based on its efficiency and convenience for respondents, enabling them to complete the survey promptly [16]. The survey was deployed in Indonesia, employing a Likert scale ranging from 1 (strongly disagree) to 6 (strongly agree). The use of a six-point Likert scale was preferred for its potential for enhanced reliability and construct validity compared to a five-point Likert scale [38].

All items were originally in English and were later translated into Bahasa Indonesia following the translation-back-translation procedure. The questionnaire's first section collected job and demographic information, serving as a filter to ensure responses came from participants with a relevant background. The questionnaire also included a metadata component guide for added clarity. To validate the translated items and the guide, one manager and one senior manager from the e-commerce industry were selected to review them.

5.4. Population and Sampling

The population of this study is defined as follows:

- Employees working at the enterprise organization in Indonesia.
- Employees who have access to BI systems with metadata are available in some form.
- Employees whose day-to-day tasks include finding and analyzing data, such as business intelligence specialists, data scientists, data analysts, or data managers.

In defining the minimum sample size, this study refers to the guidelines of Gefen et al. [40], in which, based on the use of structural equation modeling with partial least squares applied (PLS-SEM), the required minimal sample size is at least 10 times the number of items in the construct. This study has 28 items; thus, the required minimal sample size is 280 samples.

Data collection for this study was conducted from 1 December 2023 to 30 December 2023. Microsoft Forms was used to generate the questionnaire and was distributed over various social media such as WhatsApp and LinkedIn, as well as direct email to corresponding data-related personnel. The questionnaire was also spread among data community members for greater reach. Out of 521 responses collected, 27 were out of criteria. These responses either do not have metadata or BI systems implemented in their company or whose jobs are not related to finding or analyzing data. Furthermore, 176 were excluded due to duplicate entries. These duplicate entries were identified from the corresponding email or phone number during the submission. The result was 318 valid responses used during the analysis process.

5.5. Measurement Model

Partial least square structural equation modeling (PLS-SEM) was used for both data analysis and hypothesized model testing. The statistical data analysis performed through SEM consists of a measurement model and a structural model. The relations between the observed and unobserved variables are tested through the measurement model, while the structural model analyzes the path to identify any direct or indirect interactions between the unobserved or latent variables [41]. In hypothesis model testing, all hypothesized relationships are analyzed simultaneously.

Numerous previous studies on metadata quality and business intelligence [13, 16, 18] were also analyzed with PLS-SEM. PLS-SEM was chosen over CB-SEM since this study aims to explore or predict the relations between metadata quality and organizational effectiveness. The analysis was conducted using the Smart PLS 4.0 application.

6. Result and Analysis

6.1. Respondent Profile

Since the questionnaire of this study was distributed through an online survey, the respondents were coming from various industries as shown in Table 6. The demographic profile of the respondents was dominated by employees working in IT-related industries, with 65% of them currently employed as staff. The majority of them hold bachelor's degrees and come from SMEs with less than 500 employees within the company.

Table 6. Respondents' Demographic Profile

Category	Frequency	Percentage
Educational Level		
High School	58	18%
Bachelor	232	73%
Masters	13	4%
Others	15	5%
Job Level		
Staff	207	65%
Supervisor	46	14%
Manager	52	16%
Director/C Level	13	4%
Industry		
IT Software, Hardware, and Services	51	16%
Manufacturing	48	15%
Retail and Commerce	36	11%
Consumer Goods Industry	31	10%
Finance	29	9%
Education	29	9%
Construction & Real Estate	22	7%
Corporate Services	18	6%
Energy & Mining	17	5%
Media Communications and Design	15	5%
Transportation & Logistics	12	4%
Others	10	3%
Company Size		
≤ 500 employees	136	43%
501-1,000 employees	83	26%
1,001-5,000 employees	70	22%
5,001-10,000 employees	11	3%
>10,000 employees	18	6%

6.2. Measurement Model Assessment

The first assessment is to test the validity and reliability of the proposed research model. The construct loading factor value on the latent variable is used to check the convergent validity. According to Hair et al. [42] study, the expected Average Variance Extracted (AVE) value should be ≥ 0.5 to be considered valid. Table 7 demonstrates that the factor loadings of all items ranged from 0.692 to 0.894, thus exceeding the recommended value. To verify whether each construct is divergent between one another, the discriminant validity test is performed. According to Henseler et al. [43] evaluating the discriminant property should be done carefully to avoid any hindsight from the limitations in a factor model setting. Therefore, the Fornell-Larcker criterion [44] and the HTMT-based cross-loading factor are used to evaluate the discriminant validity of this study. The Fornell-Larcker criterion defines discriminant validity to be achieved when the square root of the AVE of each construct is higher than the construct's highest correlation with any other constructs [44]. The Fornell-Larcker criterion results are shown in Table 8, which shows the requirement has been fulfilled. On the other hand, the HTMT criterion requires the value of distinct constructs to be lower than 0.9 [42, 43]. This criterion is also fulfilled based on the results shown in Table 9; thus, the model is considered valid.

For reliability testing, Cronbach's Alpha and Composite Reliability Tests are deployed. Both the Cronbach's Alpha and the Composite Reliability value should be greater than 0.7 for a construct to be considered reliable [40, 42]. As displayed in Table 7, Cronbach's Alpha value ranged from 0.780 to 0.911, while the Composite Reliability value ranged from 0.872 to 0.933. Therefore, it can be concluded that all items are statistically reliable.

Table 7. Properties of the Measurement Model

Constructs	Items	Mean	Standard Deviation	Loading	Cronbach's Alpha	Composite Reliability	AVE
Descriptive Metadata	DM1	4.840	1.044	0.862	0.834	0.901	0.751
	DM2	4.805	0.997	0.888			
	DM3	4.921	0.976	0.849			
Administrative Metadata	AM1	4.840	0.930	0.862	0.871	0.912	0.722
	AM2	4.862	0.974	0.845			
	AM3	4.899	0.943	0.844			
	AM4	4.950	0.944	0.847			
Structural Metadata	SM1	4.755	0.976	0.894	0.867	0.918	0.79
	SM2	4.752	1.027	0.886			
	SM3	4.833	1.052	0.886			
Metadata Management	MM1	4.937	0.923	0.788	0.780	0.872	0.695
	MM2	4.912	0.853	0.867			
	MM3	4.799	0.963	0.844			
BI System Effectiveness	BI1	4.903	0.879	0.693	0.904	0.92	0.536
	BI2	4.764	0.980	0.712			
	BI3	4.915	0.881	0.732			
	BI4	5.079	0.845	0.749			
	BI5	5.025	0.904	0.720			
	BI6	4.909	0.919	0.692			
	BI7	4.736	1.015	0.720			
	BI8	4.799	0.973	0.759			
	BI9	4.736	1.052	0.774			
	BI10	4.780	1.059	0.767			
Organizational Effectiveness	OE1	4.670	0.972	0.849	0.911	0.933	0.737
	OE2	4.695	1.027	0.851			
	OE3	4.811	1.047	0.893			
	OE4	4.824	0.949	0.841			
	OE5	4.947	1.022	0.859			

Table 8. Fornell-Larcker Criterion Value

Constructs	DM	AM	SM	MM	BI	OE
DM	**0.867**					
AM	0.752	**0.849**				
SM	0.720	0.779	**0.889**			
MM	0.664	0.715	0.693	**0.834**		
BI	0.623	0.695	0.669	0.706	**0.732**	
OE	0.591	0.560	0.627	0.533	0.611	**0.859**

Table 9. Heterotrait-Monotrait (HTMT) Ratio

Constructs	DM	AM	SM	MM	BI	OE
DM						
AM	0.881					
SM	0.847	0.896				
MM	0.822	0.865	0.841			
BI	0.714	0.778	0.749	0.832		
OE	0.677	0.627	0.705	0.631	0.669	

6.3. Structural Model Assessment

After validating the measurement model, the bootstrapping technique, employing a minimum sample size of 5,000 subsamples, is utilized to assess the significance and relevance of path coefficients in the structural model [45]. Subsequently, the explanatory power of the structural model is gauged by calculating the coefficient of determination (R^2) for each dependent variable. The R^2 measures the predictive accuracy of the model with recommended values of 0.75 for high, 0.50 for moderate, and 0.25 for weak levels of predictive accuracy [42]. As shown in Table 10, the research structural model can explain BI Systems Effectiveness and Organizational Effectiveness with relatively low and moderate accuracy, respectively. This implies that 59% of any changes in BI Systems Effectiveness and 48% of any changes in Organizational Effectiveness can be explained by other constructs affecting these two variables.

Table 10. Coefficient of Determination and Stone-Geisser Test

Constructs	R^2	Q^2
BI Systems Effectiveness	0.588	0.567
Organizational Effectiveness	0.479	0.408

The Stone-Geisser (Q^2) metric, used to assess the predictive relevance of the model, is also employed [45]. A Q^2 value exceeding 0.02 indicates small, 0.15 suggests moderate, and 0.35 signifies large predictive relevance of the PLS path model [42]. According to the results shown in Table 10, the research model demonstrates a large predictive relevance for both BI Systems Effectiveness and Organizational Effectiveness.

Subsequently, a multicollinearity analysis is conducted to elucidate the degree to which an independent variable varies with other independent variables [40]. There are various recommendations for the acceptable levels of variance inflation factor (VIF), used as an indicator of multicollinearity. Hair et al. [42] suggested that the VIF value is expected to be lower than 5, while James et al. [46] proposed that a VIF value between 5 and 10 indicates a moderate correlation, while a value larger than 10 is not tolerable. As seen in Table 11, the VIF values are all below 5, which means each variable is different from the other.

Table 11. Variance Inflation Factor Values and Tolerance Level

Path	VIF	Tolerance Level	Result
DM→BI	2.673	0.374	Low
DM→OE	2.685	0.372	Low
AM→BI	3.435	0.291	Low
AM→OE	3.581	0.279	Low
SM→BI	3.025	0.330	Low
SM→OE	3.100	0.322	Low
MM→BI	2.353	0.424	Low
MM→OE	2.669	0.374	Low
BI→OE	2.425	0.412	Low

6.4. Hypotheses Testing

In testing the hypotheses of this study, both direct and indirect paths of each interconnected construct were thoroughly examined using Smart PLS 4.0. The significance and relevance of the structural model were determined through bootstrapping with a minimum sample size of 5,000 subsamples and a significance level of 5%. Figure 5 displays the graphical output of each direct path, along with the R^2 value for the endogenous latent variables.

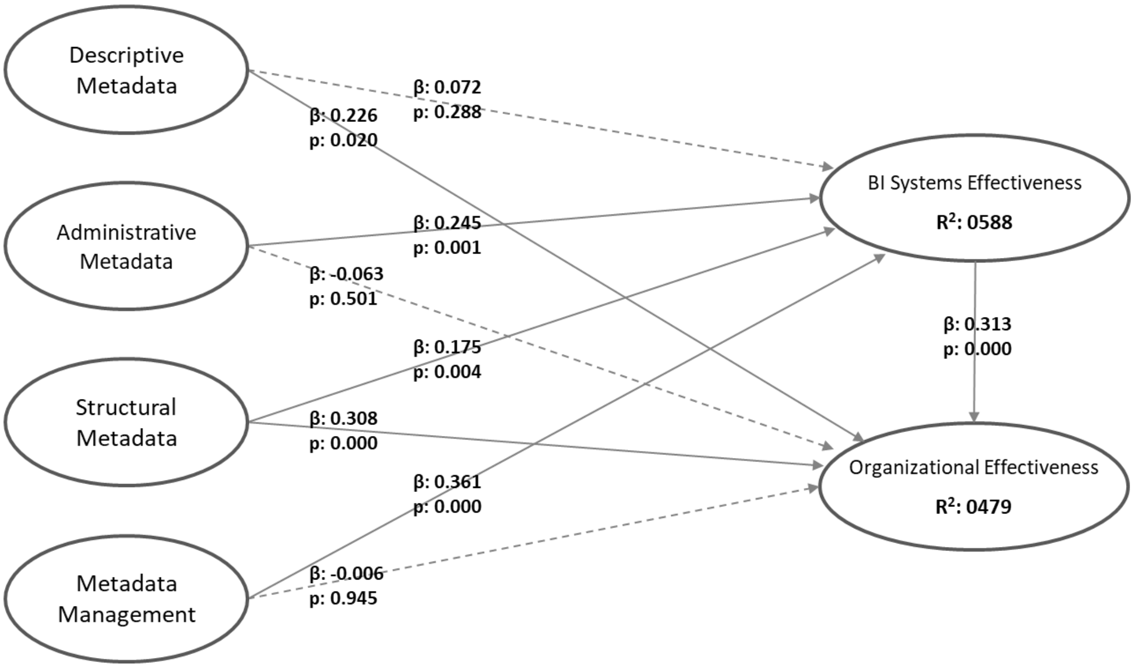

Figure 5. Research Model with Direct Geographical Result

The unmediated structural model testing result is shown in Table 12. Of the nine hypotheses, three are rejected, which are the relationship of Descriptive Metadata to BI Systems Effectiveness ($\beta = 0.072$, $t = 1.063$, $p = 0.288$), Administrative Metadata to Organizational Effectiveness ($\beta = -0.063$, $t = 0.673$, $p = 0.501$), and Metadata Management to Organizational Effectiveness ($\beta = -0.006$, $t = 0.069$, $p = 0.945$). These conclusions are made since the *p*-value of the path connecting those variables is greater than 0.05, which means the effect is not significant.

Table 12. Hypotheses Testing Result of Direct Effect

Hypothesis	Path	Coefficient	T Value	P Value	Result
H1	BI→OE	0.313	4.389	0.000	Supported
H2	DM→BI	0.072	1.063	**0.288**	**Rejected**
H3	DM→OE	0.226	2.324	0.020	Supported
H5	AM→BI	0.245	3.387	0.001	Supported
H6	AM→OE	-0.063	0.673	**0.501**	**Rejected**
H8	SM→BI	0.175	2.877	0.004	Supported
H9	SM→OE	0.308	3.686	0.000	Supported
H11	MM→BI	0.361	6.333	0.000	Supported
H12	MM→OE	-0.006	0.069	**0.945**	**Rejected**

The test results for hypotheses involving mediation effects are presented in Table 13. Among the four hypotheses related to mediating effects, one is rejected, specifically the mediation effect of BI Systems Effectiveness in the relationship between Descriptive Metadata and Organizational Effectiveness ($\beta = 0.023$, $t = 0.970$, $p = 0.332$). Additionally, it is observed that BI Systems Effectiveness fully mediates the relationships between Administrative Metadata and Organizational Effectiveness ($\beta = 0.077$, $t = 2.570$, $p = 0.010$) and Metadata Management and Organizational Effectiveness ($\beta = 0.113$, $t = 3.535$, $p = 0.000$). This conclusion is drawn as there is no significant direct path from Administrative Metadata to Organizational Effectiveness and from Metadata Management to Organizational Effectiveness. H10 ($\beta = 0.055$, $t = 2.331$, $p = 0.020$), involving the mediating effect of BI Systems Effectiveness between Structural Metadata and Organizational Effectiveness, demonstrates partial mediation, given the significant effect observed in the direct path from Structural Metadata to Organizational Effectiveness.

Table 13. Hypotheses Testing Result of Indirect Effect

Hypothesis	Path	Coefficient	T Value	P Value	Result	Mediation
H4	DM→BI→OE	0.023	0.970	**0.332**	**Rejected**	**None**
H7	AM→BI→OE	0.077	2.570	0.010	Supported	Full
H10	SM→BI→OE	0.055	2.331	0.020	Supported	Partial
H13	MM→BI→OE	0.113	3.535	0.000	Supported	Full

7. Qualitative Study

To supplement the findings from the quantitative survey data and to capture diverse perspectives in interpreting the results, a focus group discussion was conducted. The focus group is chosen as it provides more study information [47] and works with interested and knowledgeable participants [48], which aligns with the survey participants from the quantitative study. By capturing wider perspectives, the focus group also facilitates a deeper understanding of metadata implementation and its impact across various organizational settings and industries. The discussion was conducted via an online meeting and focused on the results of testing the 13 hypotheses.

7.1. Selection of Participants

To ensure that all participants had pre-existing knowledge of metadata and BI systems implementation research, we contacted those who had previously opted to be reached for further analysis. The discussion included 17 participants working as data analysts, business intelligence specialists, data engineers, data scientists, and data infrastructure implementers. These participants represented a diverse range of industries, including software and IT services, retail and commerce, finance, energy and mining, food and beverage, recreation and travel, education, and corporate services.

7.2. Focus Group Discussion Result

Each hypothesis testing result was presented consecutively to the participants. During the discussion, all participants were given time to compare the results from the quantitative study with the conditions in their own workplaces. Table 14 shows the focus group discussion results, with "a" denoting agreement and "d" denoting disagreement with the results. An asterisk (*) indicates that further arguments were presented by the participants to support their views. These additional perspectives provide invaluable qualitative information to support the overall findings of this metadata research.

Two out of the thirteen hypothesis testing results were mostly disagreed upon by the experts. These hypotheses were the significance test on the impact of descriptive metadata on BI systems effectiveness, and the impact of descriptive metadata on organizational effectiveness, considering the mediating effect of BI systems effectiveness.

Table 14. Focus Group Discussion Result

Hypothesis	Path	Result	Participant																
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
H1	BI→OE	Significant	a	a	a	a	a	a	a	a	a*	a*	a*	a*	a	a	a	a	a
H2	DM→BI	Not significant	a*	a*	a	d*	d*	d*	d*	d*	d*	d*	d	d*	a	a	a	a*	d*
H3	DM→OE	Significant	a	a	a	a	a	a	a	a*	a	a	a	a	a	a	a	a	a
H4	DM→BI→OE	Not significant	a	a	a	d*	d*	d*	d*	d*	d*	d*	d*	d*	a	a	a	a	d*
H5	AM→BI	Significant	a*	a	a	a	a	a	a	a*	a	a	a	a	a	a	a	a*	a
H6	AM→OE	Not significant	a	a	a	a	a	a	d*	a	a	a	a	a	a	a	a	a	a
H7	AM→BI→OE	Significant	a	a	a	a	a	a	a	a	a	a*	a	a	a	a	a	a	a
H8	SM→BI	Significant	a*	a*	a	a	a	a	a	a	a*	a	a*	a*	a	a	a	a	a
H9	SM→OE	Significant	a	a	a	a	a	a	a	a	a	a	a	a*	a	a	a	a	a
H10	SM→BI→OE	Significant	a	a	a	a	a	a	a*	a	a	a	a	a	a	a	a	a	a
H11	MM→BI	Significant	a*	a*	a	a	a	a	a	a	a*	a*	a	a*	a	a	a	a*	a
H12	MM→OE	Not significant	a	a*	a	a	a	a	d*	d*	a*	a	a	a	a	a	a	a	a*
H13	MM→BI→OE	Significant	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a	a

8. Discussion

This study has empirically developed and measured the metadata implementation components in determining the effectiveness of organizations' BI systems, which lead to the effectiveness of the organization. The first finding is the significant effect of BI systems effectiveness on organizational effectiveness. The positive relationship is consistent with previous studies that support the argument that BI systems effectiveness is influencing organizational effectiveness [16, 18]. Additional insights from the qualitative study, gathered through the focus group discussion, suggest that BI systems are beneficial for onboarding new team members, helping them quickly understand organizational data. This advantage also extends to employees across different functions, enabling them to familiarize themselves with data from other departments, thereby further enhancing organizational effectiveness through better data utilization.

In terms of metadata components' effect on the effectiveness of BI systems, this study found that administrative metadata, structural metadata, and metadata management have significant effects. Administrative metadata captures the technical information of data, while structural metadata helps users in understanding the relationship between the data with the business context and its origin [7, 9, 35]. Two participants from the group discussion specifically highlighted the importance of accessibility information in administrative metadata, noting that it becomes increasingly complex as an organization grows. Additionally, the participants emphasized the benefits of structural metadata for information auditing, as it provides data calculation and transformation details for metrics. These features are crucial and can only be fully leveraged through the implementation of BI systems for accessing metadata information.

Metadata management is the strongest predictor of BI systems effectiveness, with a path coefficient of 0.361. This finding is parallel to the previous study, which revealed that the navigational metadata, characteristics that define the metadata findability attributes, have significant effects on the acceptance of BI systems [16]. Metadata management is closely related to the findability and accessibility of the FAIR element [19, 37]; hence, its effectiveness is closely tied to the BI systems but does not directly affect the organizational effectiveness. This is shown by the result of this study in which the path coefficient from metadata management to organizational effectiveness is the lowest and insignificant, while, with the mediation effect of BI, the path coefficient is the highest.

Additional insights from the qualitative study also validate the importance of automation and collaborative processes in supporting the sustainability of overall metadata management, thereby enhancing the relevance of BI systems. One participant specifically highlighted the approval capability in collaborative efforts to ensure the validity of the metadata. A participant from a finance startup company emphasized that involving the business team in the collaborative process greatly helps in capturing relevant business terms, which can grow sporadically in a startup environment. BI systems are seen as an integral part of metadata management, with their impact on organizational effectiveness being mediated by the system itself.

The hypothesis on the influence of descriptive metadata on BI systems is, however, rejected. Descriptive metadata provides general information about data [7, 9] and this information is foundational to be part of not only the BI systems but also in essentially every data-related application, such as relational database and data warehouse [35, 37]. However, most focus group participants disagreed with this result, including the finding on the insignificant effect of BI systems as a mediator. They considered the redundancy of general information, such as titles, descriptions, and data formats, presented in both BI systems and databases as necessary to avoid the need for switching between different applications. Moreover, access to databases to obtain basic information might not be available to non-technical users, which would hinder their ability to retrieve accurate data.

On the influence of metadata components on organizational effectiveness, descriptive metadata and structural metadata are confirmed to have a significant effect. One can conclude that the foundational nature of descriptive metadata, which is the general information of the data, as well as the business context and linkage information provided by the structural metadata enables an organization to make sense of the data definition from business terms across all functional silos [21, 35]. Thus, increasing the effectiveness of an organization. This argument was also supported by the focus group participants. Structural metadata information also benefited from the BI System implementation since all the data relationship information will be easier to understand if presented in an interface of some sort.

The effect of administrative metadata on organizational effectiveness has the same nature as metadata management. The administrative metadata, which holds the technical information of data such as data lineage [7, 9], is confirmed to have significant effects on the BI systems' effectiveness and is aligned with the findings of previous research [16]. But instead of directly affecting the organizational effectiveness, its effect is mediated through the BI systems implementation. This is understandable as the lineage information should be presented visually through a BI system's interface for it to bring benefit, similar to the nature of the search mechanism in metadata management, which needs to be implemented in a BI system. Specifically, focus group participants highlighted data lineage and modification tracking information as essential components for troubleshooting when faulty data is presented and root cause identification is needed.

9. Conclusion

This study significantly enhances our understanding of metadata components and their impact on Business Intelligence system effectiveness and organizational effectiveness. Recognizing the interplay among these factors is crucial for organizations aiming to achieve FAIR data and informed decision-making. A systematic literature review and a preliminary study have helped to understand the landscape of metadata categorization. Then through empirical study, we reveal that administrative metadata, structural metadata, and metadata management directly influence Business Intelligence systems effectiveness, subsequently affecting organizational effectiveness. Notably, a Business Intelligence system mediates the relationships between administrative metadata, structural metadata, and metadata management. These insights guide organizations in prioritizing metadata implementation during early adoption stages. Additional insights from the focus group discussion also revealed important metadata roles in different industries. The organization also benefits from the findings of this research in terms of prioritizing the metadata that should be implemented during the early stages of metadata adoption. The preliminary study demonstrated that the most anticipated metadata capability is the capacity to identify the source of data from an application; however, the implementation is relatively subpar. This capability is essential for assisting both data and business users in identifying the fundamental data source of a specific User Interface application, thereby enhancing the data's findability. Although it is difficult to prioritize this capability, it will be advantageous for both organizational effectiveness and Business Intelligence systems, as it is a component of the structural metadata. Overall, this research offers actionable knowledge for optimizing metadata strategies and fostering data-driven decision-making processes in organizations.

9.1. Limitations and Future Directions

While this study provides valuable insights into the impact of metadata components on BI systems and organizational effectiveness, it acknowledges several limitations warranting future exploration. Grouping metadata components into broad categories risks overgeneralization, suggesting a need for more nuanced evaluations of individual components. For example, separating indicators for table descriptions and data ranges within the descriptive metadata category could enhance the precision and understanding of each component's unique contribution.

Moreover, the study's reliance on cross-sectional surveys hinders the ability to establish direct causality between metadata components and efficiency gains. Exploring experimental designs and constructing public data warehouses for broader participation present challenges but offer avenues for future research. While the survey captured responses from diverse industries, it's crucial to recognize that different sectors may have unique data requirements. The additional qualitative study through the focus group discussion, which involved participants from various industries, aims to minimize this limitation. However, studies focusing on specific industries or conducting cross-industry comparisons can still elucidate industry-specific metadata utilization patterns, enhancing the study's applicability and understanding of varied organizational contexts.

10. Declarations

10.1. Author Contributions

Conceptualization, T.O. and E.W.; methodology, T.O.; software, E.W.; validation, T.O. and E.W.; formal analysis, T.O.; investigation, T.O.; resources, T.O.; data curation, E.W.; writing—original draft preparation, T.O and E.W.; writing—review and editing, T.O.; visualization, E.W.; supervision, T.O.; project administration, T.O.; funding acquisition, T.O. All authors have read and agreed to the published version of the manuscript.

10.2. Data Availability Statement

The data presented in this study are available in the article.

10.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

10.4. Institutional Review Board Statement

Not applicable.

10.5. Informed Consent Statement

Not applicable.

10.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

11. References

[1] Peyton, L. (2009). Common Warehouse Metamodel. Encyclopedia of Database Systems. Springer, Boston, United States. doi:10.1007/978-0-387-39940-9_900.

[2] Labadie, C., Eurich, M., & Legner, C. (2020). Empowering data consumers to work with data: Data documentation for the enterprise context. Proceedings of the 15th International Conference on Business Information Systems 2020 "Developments, Opportunities and Challenges of Digitization", WIRTSCHAFTSINFORMATIK 2020, 1428-1442. doi:10.30844/wi_2020_m7.

[3] Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... & Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data, 3(1), 1-9. doi:10.1038/sdata.2016.18.

[4] Inmon, W. H., O'Neil, B. K., & Fryman, L. (2007). Business Metadata: Capturing Enterprise Knowledge. Morgan Kaufmann Publishers, Massachusetts, United States. doi:10.1016/B978-0-12-373726-7.X5001-9.

[5] Roszkiewicz, R. (2010). Enterprise metadata management: How consolidation simplifies control. Journal of Digital Asset Management, 6(5), 291–297. doi:10.1057/dam.2010.32.

[6] Eichler, R., Giebler, C., Gröger, C., Schwarz, H., & Mitschang, B. (2020). Handle - a generic metadata model for data lakes. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 12393 LNCS, 73–88. doi:10.1007/978-3-030-59065-9_7.

[7] Riley, J. (2017). Understanding metadata. NISO Primer Series: National Information Standards Organization, Washington D.C., United States Available online: https://groups.niso.org/higherlogic/ws/public/download/17446/Understanding%20Metadata.pdf (accessed on May 2024).

[8] Shanmugam, S., & Seshadri, G. (2016). Aspects of Data Cataloguing for Enterprise Data Platforms. Proceedings - 2nd IEEE International Conference on Big Data Security on Cloud, IEEE Big Data Security 2016, 2nd IEEE International Conference on High Performance and Smart Computing, IEEE HPSC 2016 and IEEE International Conference on Intelligent Data and Security, IEEE IDS 2016, 134–139. doi:10.1109/BigDataSecurity-HPSC-IDS.2016.52.

[9] Ehrlinger, L., Schrott, J., Melichar, M., Kirchmayr, N., & Wöß, W. (2021). Data Catalogs: A Systematic Literature Review and Guidelines to Implementation. Communications in Computer and Information Science: Vol. 1479 CCIS, 148–158. doi:10.1007/978-3-030-87101-7_15.

[10] Gabriel, R., Hoppe, T., & Pastwa, A. (2010). Classification of metadata categories in data warehousing - A generic approach. In 16th Americas Conference on Information Systems 2010, AMCIS 2010, 2, 1539–1549.

[11] Quimbert, E., Jeffery, K., Martens, C., Martin, P., & Zhao, Z. (2020). Data cataloguing. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 12003, 140–161. doi:10.1007/978-3-030-52829-4_8.

[12] Kiron, D., Shockley, R., Kruschwitz, N., Finch, G., & Haydock, M. (2012). Analytics: The widening divide. MIT Sloan Management Review, 53(2), 3–20.

[13] Foshay, N., Taylor, A., & Mukherjee, A. (2014). Winning the Hearts and Minds of Business Intelligence Users: The Role of Metadata. Information Systems Management, 31(2), 167–180. doi:10.1080/10580530.2014.890444.

[14] Ehrenmann, M., Pieringer, R., & Stockinger, K. (2012). Is There a Cure-All for Business Analytics? Business Intelligence Journal, 17(3), 28–39.

[15] Hassani, A. L., & Mansouri, K. (2017). Towards a Model of a Scalable Middleware Architecture Based on Cloud Computing, Application Reliable Integration and E-Learning Platforms in Academic Information Systems. Transactions on Machine Learning and Artificial Intelligence, 5(4), 651-664. doi:10.14738/tmlai.54.3329.

[16] Masa'Deh, R., Obeidat, Z., Maqableh, M., & Shah, M. (2021). The Impact of Business Intelligence Systems on an Organization's Effectiveness: The Role of Metadata Quality from a Developing County's View. International Journal of Hospitality and Tourism Administration, 22(1), 64–84. doi:10.1080/15256480.2018.1547239.

[17] Kara, H. (2023). Research and Evaluation for Busy Students and Practitioners. Policy Press, Bristol, England. doi:10.56687/9781447366263.

[18] Arefin, M. S., Hoque, M. R., & Bao, Y. (2015). The impact of business intelligence on organization's effectiveness: An empirical study. Journal of Systems and Information Technology, 17(3), 263–285. doi:10.1108/JSIT-09-2014-0067.

[19] Labadie, C., Legner, C., Eurich, M., & Fadler, M. (2020). FAIR Enough? Enhancing the Usage of Enterprise Data with Data Catalogs. In Proceedings - 2020 IEEE 22nd Conference on Business Informatics, 201–210. doi:10.1109/CBI49978.2020.00029.

[20] Chelmis, C., Zhao, J., Sorathia, V., Agarwal, S., & Prasanna, V. (2013). Toward an automatic metadata management framework for smart oil fields. SPE Economics and Management, 5(1), 33–43. doi:10.2118/153271-PA.

[21] Dela Cruz, N., Schiefelbein, P., Anderson, K., Hallock, P., & Barden, D. (2010). ORM and MDM/MMS: Integration in an enterprise level conceptual data model. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 6428 LNCS, 457–463. doi:10.1007/978-3-642-16961-8_67.

[22] Dzyubanenko, A. A., & Rabin, A. V. (2022). Hybrid client- server implementation and microservice architecture of automatic documentation analysis software. CEUR Workshop Proceedings, 3091, 93–101. doi:10.47813/dnit-mip5/2022-3091-93-101.

[23] Vieira, E. S., & Gomes, J. A. N. F. (2009). A comparison of Scopus and Web of science for a typical university. Scientometrics, 81(2), 587–600. doi:10.1007/s11192-009-2178-0.

[24] Chen, L., Babar, M. A., & Zhang, H. (2010). Towards an Evidence-Based Understanding of Electronic Data Sources. doi:10.14236/ewic/ease2010.17.

[25] Carrera-Rivera, A., Ochoa, W., Larrinaga, F., & Lasa, G. (2022). How-to conduct a systematic literature review: A quick guide for computer science research. MethodsX, 9. doi:10.1016/j.mex.2022.101895.

[26] Zaidi, E., De Simoni, G., Edjlali, R., & Duncan, A. D. (2017). Data Catalogs Are the New Black in Data Management and Analytics. Gartner, December, 1–16.

[27] Qi, X. (2021). Research on Enterprise Data Governance Based on Knowledge Map. Advances in Intelligent Systems and Computing, 1283, 581–586. doi:10.1007/978-3-030-62746-1_85.

[28] Yu, Q. C. (2010). Metadata integration architecture in enterprise data warehouse system. 2nd International Conference on Information Science and Engineering, ICISE2010 - Proceedings, 340–343. doi:10.1109/ICISE.2010.5690320.

[29] Schutz, C., & Schrefl, M. (2014). Customization of Domain-Specific Reference Models for Data Warehouses. IEEE 18th international Enterprise Distributed object computing conference, 61–70. doi:10.1109/EDOC.2014.18.

[30] Fujita, Y., Naono, K., & Hanai, T. (2012). Proposal and evaluation of metadata management method for eDiscovery. Annual SRII Global Conference, SRII, 778–786. doi:10.1109/SRII.2012.88.

[31] Yan, Y., & McLane, T. (2012). Metadata management and revision history tracking for spatial data and GIS map figures. ACM International Conference Proceeding Series, 1–2. doi:10.1145/2345316.2345357.

[32] van Helvoirt, S., & Weigand, H. (2015). Operationalizing data governance via multi-level metadata management. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9373, 160–172. doi:10.1007/978-3-319-25013-7_13.

[33] Seng, J. L., & Wong, Z. (2012). An intelligent XML-based multidimensional data cube exchange. Expert Systems with Applications, 39(8), 7371–7390. doi:10.1016/j.eswa.2012.01.069.

[34] Vnuk, L., Koronios, A., & Gao, J. (2011). Enterprise metadata management: Conceptions, issues and capabilities. ACIS 2011 Proceedings - 22nd Australasian Conference on Information Systems, 83.

[35] Petrik, D., Untermann, A., & Baars, H. (2024). Functional Requirements for Enterprise Data Catalogs: A Systematic Literature Review. Lecture Notes in Business Information Processing: Vol. 500 LNBIP, 3–18. doi:10.1007/978-3-031-53227-6_1.

[36] Nizamis, A., Ioannidis, D., Gkonis, P., & Trakadas, P. (2024). Introducing an Enhanced Metadata Broker for Manufacturing Data Spaces. ACM International Conference Proceeding Series, 37–40. doi:10.1145/3685651.3686699.

[37] Eichler, R., Gröger, C., Hoos, E., Stach, C., Schwarz, H., & Mitschang, B. (2023). Introducing the enterprise data marketplace: a platform for democratizing company data. Journal of Big Data, 10(1), 173. doi:10.1186/s40537-023-00843-z.

[38] Abdul. (2010). Quality of Psychology Test Between Likert Scale 5 and 6 Points. Journal of Social Sciences, 6(3), 399–403. doi:10.3844/jssp.2010.399.403.

[39] Turban, E., Sharda, R., & Delen, D. (2010). Decision Support and Business Intelligence Systems. Pearson Education India, Chennai, India.

[40] Gefen, D., Straub, D., & Boudreau, M.-C. (2000). Structural Equation Modeling and Regression: Guidelines for Research Practice. Communications of the Association for Information Systems, 4. doi:10.17705/1cais.00407.

[41] Kline, R. B. (2007). Principles and practices of structural equation modelling 4th edition. Methodology in the social sciences, The Guilford Press, New York, United States.

[42] Hair, J., Hollingsworth, C. L., Randolph, A. B., & Chong, A. Y. L. (2017). An updated and expanded assessment of PLS-SEM in information systems research. Industrial Management and Data Systems, 117(3), 442–458. doi:10.1108/IMDS-04-2016-0130.

[43] Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. Journal of the Academy of Marketing Science, 43(1), 115–135. doi:10.1007/s11747-014-0403-8.

[44] Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. Journal of Marketing Research, 18(1), 39–50. doi:10.1177/002224378101800104.

[45] Hair Jr, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). Partial least squares structural equation modeling (PLS-SEM) using R: A workbook. Springer, Cham, Switzerland. doi:10.1007/978-3-030-80519-7.

[46] James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning: With Applications in R, Springer Texts in Statistics, 127–173. Springer, New York, United States. doi:10.1007/978-1-4614-7138-7_4

[47] Pushkarenko, K., Cavell, M., Gosse, N., & Michalovic, E. (2023). Physical literacy and the participant perspective: Exploring the value of physical literacy according to individuals experiencing disability through composite narratives. Journal of Exercise Science and Fitness, 21(3), 237–245. doi:10.1016/j.jesf.2023.03.001.

[48] López Secanell, I. (2023). Analysis of the perception of students and teachers about physical education proposals around contemporary art: a case study. Retos, 49, 260–269. doi:10.47197/RETOS.V49.98182.

Appendix I: Questionnaires

Construct	Item in English
Descriptive Metadata (DM)	DM1: The data description provided is complete enough to help me understand the meaning of data. DM2: Generally, I find the data description to be accurate. DM3: Data descriptions are presented in a format that is clear and understandable.
Administrative Metadata (AM)	AM1: Generally, the administrative information of the data provided is accurate. AM2: I get all the information I need to understand the data administration. AM3: The administrative information provided is dependable. AM4: Administrative information is presented in a format that is clear and understandable.
Structural Metadata (SM)	SM1: The data structure information provided is accurate. SM2: I get all the information I need to understand the structure of data. SM3: The data structure information provided is dependable.
Metadata management (MM)	MM1: I have access to a good data search facility. MM2: The BI application provided is effective in helping me to locate the data. MM3: Generally, it is easy for me to find the data I need, even if I have not used the data before.
BI Systems Effectiveness (BI)	BI1: BI system improves coordination with business partners/other departments. BI2: BI system reduces the cost of transactions with business partners/other departments. BI3: BI system improves responsiveness to/from other departments. BI4: BI system improves efficiency of internal processes. BI5: BI system increases staff productivity. BI6: BI system reduces the cost of effective decision-making. BI7: BI system reduces operational cost. BI8: BI system reduces customer return handling costs. BI9: BI system reduces marketing costs. BI10: BI system reduces time-to-market products/services.
Organizational Effectiveness (OE)	OE1: Compared with key competitors, our company is more successful. OE2: Compared with key competitors, our company has a greater market share. OE3: Compared with key competitors, our company is growing faster. OE4: Compared with key competitors, our company is more profitable. OE5: Compared with key competitors, our company is more innovative.