Communication Pattern in social media Metadata with FAIRdata Perspective

Author:

Agus Pamuji[1]

Rina Rindanah[2]

M. Reyhan Fadilah[3]

Affiliation:

[1,2] Universitas Islam Negeri Siber Syekh Nurjati, Indonesia

Corresponding author:

agus.pamuji@syekhnurjati.ac.id[1]

rinarindanah@syekhnurjati.ac.id[2]

reyhanmuhamamd21@gmail.com[3]

The rise of electronic devices has had a profound impact on the ways in which digital communication tools are utilized within Islamic higher education. This change emphasizes the growing importance of adhering to the FAIR data principles – Findable, Accessible, Interoperable, and Reusable. Despite a significant body of research on these principles, there is a noticeable lack of emphasis on the formal construction of metadata to facilitate seamless information exchange among users and devices. Recognizing this gap, our research is focused on extensively exploring the complexities of metadata usage within Islamic higher education, particularly in the context of cyber communication. Our aim is to conduct a thorough analysis that not only evaluates the current knowledge base but also provides practical insights into implementing metadata to meet the needs of researchers, practitioners, and students engaged in scholarly communication. Additionally, we aspire to develop a comprehensive report that offers an in-depth overview of metadata.

Keywords: Pattern Communication, Metadata, Data-Driven, FAIRdata, Interoperability.

Maraknya perangkat elektronik berdampak besar pada cara penggunaan alat komunikasi digital dalam pendidikan tinggi Islam. Perubahan ini menekankan semakin pentingnya mematuhi prinsip-prinsip data FAIR – Dapat Ditemukan, Dapat Diakses, Dapat Dioperasikan, dan Dapat Digunakan Kembali. Meskipun terdapat banyak penelitian mengenai prinsip-prinsip ini, terdapat kurangnya penekanan pada konstruksi formal metadata untuk memfasilitasi pertukaran informasi yang lancar antara pengguna dan perangkat. Menyadari kesenjangan ini, penelitian kami difokuskan untuk mengeksplorasi secara ekstensif kompleksitas penggunaan metadata dalam pendidikan tinggi Islam, khususnya dalam konteks komunikasi siber. Tujuan kami adalah melakukan analisis menyeluruh yang tidak hanya mengevaluasi basis pengetahuan saat ini namun juga memberikan wawasan praktis dalam penerapan metadata untuk memenuhi kebutuhan peneliti, praktisi, dan mahasiswa yang terlibat dalam komunikasi ilmiah. Selain itu, kami bercita-cita untuk mengembangkan laporan komprehensif yang menawarkan gambaran mendalam tentang metadata

Kata Kunci: Pola Komunikasi, Metadata, Berbasis Data, FAIRdata, Interoperabilitas

INTRODUCTION

Metadata is an essential component in the effective management and utilization of information, especially within systems like Content Management Systems (CMS) or Content Delivery Portals. It enriches modular or document-based content by offering additional information, thereby facilitating the retrieval and reuse of content within CMS or in external processes during content delivery._Metadata enables more efficient and reliable information usage and provision, contributing to the overall effectiveness of these systems.

To prepare for requirement, metadata have been created by web document which the approaching has define sistematically. The proposed mechanism is designed to create RDF Semantic Web schemas from a collection of web documents by leveraging them as semantic metadata. This process involves a comprehensive examination of both the structured and unstructured content within these documents. The approach involves treating the semi-structured web documents as resource objects with interconnected relationships within the RDF graph, requiring an in-depth analysis of hyperlinks, basic annotations, and keywords present in the documents. By adhering to the proposed mechanism and the specified rules in the paper, a corresponding RDF schema will be automatically generated. The anticipated result of this approach is the seamless translation of semantic metadata from sets of documents on the Web, thus eliminating the requirement for manual editing. This advancement could open up possibilities for conducting semantic operations on the Web, including semantic queries and searches in the future.

The advancement of embedded devices with camera mobility has significantly simplified the process of creating, modifying, and sharing digital content across various social media platforms [1]. However, with this ease of sharing comes the heightened risk of spreading misinformation or disinformation, which can have detrimental effects on society. This is particularly concerning as such content can be rapidly disseminated on Online Social Networks (OSN) without undergoing proper verification for integrity and authenticity. Given the pervasive nature of this issue, it has become increasingly crucial to track the origin and nature of information being circulated on social networks [2]. A comprehensive study has been conducted to delve into the various types of metadata associated with digital images and to assess the multitude of attributes that are susceptible to compromising the authentication of the source. This vulnerability is further exacerbated by the widespread availability of online tools and mobile apps,

[1] Woojin Jung et al., "Suicidality Detection on Social Media Using Metadata and Text Feature Extraction and Machine Learning," *Archives of Suicide Research* 27, no. 1 (2023): 13–28, https://doi.org/10.1080/13811118.2021.1955783.

[2] Diogo Pacheco et al., "Uncovering Coordinated Networks on Social Media: Methods and Case Studies," *Proceedings of the International AAAI Conference on Web and Social Media* 15, no. 1 (2021): 455–466.

which can be utilized to manipulate and falsify such metadata. In conclusion, there is an urgent need to prioritize the preservation of metadata through the implementation of watermarking techniques. This is essential not only for the protection of crucial information in digital forensic investigations but also for combating the proliferation of unverified and potentially harmful content across social media platforms.

The copious volume of data associated with users, posts, and user interactions, commonly known as metadata, plays a pivotal role in shaping and enriching the services provided by online social networks (OSNs). This information is not simply an additional layer of detail but serves as the foundation for users to assess the authenticity and credibility of accounts within the network. Furthermore, the substantial body of previous research focused on combatting social spam heavily relies on the utilization of account metadata for effective detection and prevention. Metadata encompasses a wide range of data within OSNs, including details about users, post content, and the various interactions users engage in on the platform. For example, platforms like Twitter provide data on user mentions, frequency of message re-tweets, document upload times, and user interactions, all of which contribute to the metadata pool [3]. This information is not only vital for comprehending user behavior and trends, but also forms the basis for evaluating the legitimacy of user accounts. The impact and significance of metadata extend beyond mere user-related data points. It plays a pivotal role in developing and implementing effective strategies to combat social spam, thereby safeguarding the integrity of the platform.

The intricate details and patterns inherent in metadata are instrumental in identifying and addressing malicious activities, ultimately contributing to a safer and more reliable user experience. In addition to its role in assessing user credibility and detecting spam, metadata provides valuable insights for understanding user engagement and behavior patterns. Platforms like Twitter have leveraged this data to optimize user experience and security measures. Understanding and leveraging metadata is essential for enhancing the functionality and reliability of online social networks. The reliance on account metadata for in-depth analysis and research on combating social spam underscores its indispensable nature in the realm of online social networks. Its multifaceted role as a pillar of user credibility assessment, spam detection, and user behavior analysis underscores the pivotal significance of metadata in driving the efficiency and security of OSNs.

Studying urban areas through the analysis of data derived from Location Based Social Networks (LBSNs) has gained significant traction as a

[3] Atheer S. Alhassun and Murad A. Rassam, "A Combined Text-Based and Metadata-Based Deep-Learning Framework for the Detection of Spam Accounts on the Social Media Platform Twitter," *Processes* 10, no. 3 (2022).

pragmatic research approach. Nevertheless, utilizing this data presents challenges and has instigated discussions regarding its reliability. This article delves into the complexities, opportunities, as well as constraints and predispositions involved in the collection and application of LBSN data sourced from platforms like Foursquare, Twitter, Google Places, Instagram, and Airbnb for urban research. It also highlights the most recent study employing LBSN data to comprehend urban dynamics. The paper proposes a method for obtaining, selecting, categorizing, and analyzing LBSN data, and identifies key research themes based on the data components offered by these platforms. The primary contribution of this study lies in furnishing a comprehensive and intricate framework for investigating urban phenomena using LBSN data [4].

There is approaching to explain the metadata as part of FAIRdata, and essentially modeling community standards should be provided. Identifying the degree to which datasets comply with the findable, accessible, interoperable, and reusable (FAIR) principles presents challenges due to the highly specific requirements for metadata outlined in the FAIR Guiding Principles [5]. Notably, the principles stipulate that metadata must be thorough and conform to community standards relevant to the respective field. It is imperative for each scientific community to have the capability to establish their own machine-readable metadata templates that encompass these comprehensive, discipline-specific elements. Our examination of this template-centric approach has been within the framework of two software systems. The first is the CEDAR Workbench, which enables researchers to generate new metadata, and the second is the FAIRware Workbench, designed to evaluate the conformity of stored datasets' metadata to community standards [6]. The adoption of metadata templates as pivotal elements in a suite of tools for managing online datasets offers a host of advantages [7]. These templates function as a communal yardstick for defining FAIR data and encapsulate this perspective in a format conducive to dissemination across a variety of software applications to foster data governance and sharing [8].

---

[4] Mark A. Musen et al., "Modeling Community Standards for Metadata as Templates Makes Data FAIR," *Scientific Data* 9, no. 1 (2022): 1–15.

[5] Michelle L. Heacock et al., "Enhancing Data Integration, Interoperability, and Reuse to Address Complex and Emerging Environmental Health Problems," *Environmental Science and Technology* 56, no. 12 (2022): 7544–7552.

[6] Adnan Ali et al., "Real-Time Spammers Detection Based on Metadata Features with Machine Learning," *Intelligent Automation and Soft Computing* 38, no. 3 (2023): 241–258.

[7] Guilan Zhang et al., "Relationship between the Metadata and Relevance Criteria of Scientific Data," *Data Science Journal* 20, no. 1 (2021): 1–21.

[8] Hollie C. White, Leisa Gibbons, and Eileen Horansky, "Semantic Metadata as Meaning Making: Examining #hashtags and Collection Level Metadata," in *Proceedings of the International Conference on Dublin Core and Metadata Applications*, vol. 2, 2019, 37–41.

Generaly the concept of metadata communication involves the transmission and interpretation of descriptive data that provides information about other data. This fundamental concept plays a crucial role in various domains such as telecommunications, computer networks, and information systems. Metadata serves as a vital mechanism for comprehending, organizing, and overseeing data across different platforms and contexts. Its primary function is to enable efficient data retrieval, enhance data interoperability, support data integration, and facilitate accurate data analysis and interpretation [9].

The effective functioning of contemporary digital ecosystems relies heavily on metadata communication, which facilitates the precise description, interpretation, and utilization of data across various platforms and applications. It serves to improve data visibility, accessibility, and usability while providing essential support for activities such as data integration, analytics, and decision-making processes. As digital technologies continue to advance, metadata communication remains vital for ensuring streamlined data management, optimizing resource utilization, and fostering innovation across different industries.

Within Islamic higher education, the dissemination of metadata serves a pivotal function across various critical domains including course administration, research data organization, library indexing, and scholarly correspondence. The employment of metadata streamlines the efficient and proficient transfer of information, which is imperative for the progression of knowledge and scholarly endeavors within the framework of Islamic higher education. First of all, curriculum and course management, It's important to recognize the significance of metadata tags in describing academic programs, courses, and instructional materials. These tags offer vital information about content, prerequisites, and outcomes, which plays a crucial role in effective curriculum planning, course registration, and academic advising. Secondly, Research and Scholarly Communication. The text emphasizes the importance of metadata for organizing and sharing research output in Islamic higher education institutions [10]. It highlights the role of metadata in providing details about authorship, publication dates, abstracts, keywords, and citations, which facilitates the efficient discovery, access, and citation of scholarly works [11]. Finaly, library and information

[9] Brienna N. Rutherford et al., "#TurntTrending: A Systematic Review of Substance Use Portrayals on Social Media Platforms," *Addiction* 118, no. 2 (2023): 206–217.
[10] Valentina Beretta et al., "A User-Centric Metadata Model to Foster Sharing and Reuse of Multidisciplinary Datasets in Environmental and Life Sciences," *Computers and Geosciences* 154, no. July 2019 (2021): 1–10.
[11] Ariel A. Hippen and Casey S. Greene, "Expanding and Remixing the Metadata Landscape," *Trends in Cancer* 7, no. 4 (2021): 276–278, https://doi.org/10.1016/j.trecan.2020.10.011.

management. There are cataloging and information retrieval systems, providing details about library holdings and facilitating efficient search and retrieval of resources by students, faculty, and researchers across various disciplines related to Islamic studies.

Communication pattern was available for determining the metadata in Islamic Higher Education. Initially, traditional and oral communication, this focus demonstrates the cultural importance placed on interpersonal connections, deference to authority, and the passing down of knowledge through direct interaction between educators (learned individuals) and learners (followers). After that, modern Islamic institutions of higher learning focus on interdisciplinary methods and a worldwide viewpoint, resulting in varied communication styles. Furthermore, this encompasses joint research endeavors, global conferences, academic interactions was indicated to interdisciplinary and Global Perspective.

Metadata is an essential component that significantly contributes to the efficient and reliable utilization and dissemination of information, particularly in Pattern Communication. This additional contextual information provided by metadata not only enriches modular or document-based content but also aids in streamlining interoperability processes. Product metadata and information metadata are two distinct categories of metadata. Product metadata provides specific details about how a topic is related to a product, including the specific product it relates to, its attributes, and where it fits within the product ecosystem. This can include information such as product models, versions, and associated components. On the other hand, information metadata gives details about the topic itself, such as its creator, creation date, context, and the type of information it contains, such as task details, explanations, or safety advice. This type of metadata helps users understand the content, context, and relevance of the information they are consuming.

The concept of metadata pertains to information about data itself and plays a crucial role in comprehending the abundant information it encapsulates. Within documents, metadata comprises data structures such as dataset names, relationships, and fields. Commonly encompassing elements such as title, authors, time, abstract, and keywords, metadata is a fundamental aspect acknowledged by scholars. It is widely accepted that the role of metadata extends beyond providing access points for information, contributing significantly to comprehension and the assessment of relevance. The presentation and content of metadata have a profound impact on users' evaluations and satisfaction. Consequently, scholars have undertaken numerous inquiries to discern the influence of metadata on relevance assessment, including experimentation with dynamic abstracts.

In a research study conducted in 2010, Panos Balatsoukas, a prominent researcher in the field, discovered some valuable insights regarding participants' preferences for metadata. Balatsoukas found that participants expressed a strong preference for metadata that was not only easily comprehensible but also well-structured into distinct categories. To understand users' concerns related to metadata, Balatsoukas utilized an

advanced eye-tracking device to measure the participants' level of comprehension of the metadata. The results of the study indicated that different levels of cognitive effort led to varying assessments of the relevance of the metadata. The primary components of the metadata that were evaluated included the title, abstract, URL, and other relevant elements, highlighting the importance of these aspects in users' perceptions and understanding.

The role of data governance in ensuring data quality and compliance. Data governance is closely intertwined with metadata management as it establishes policies and processes for data management, ensuring that data is accurate, consistent, and reliable [12]. Effective data governance aligns with the objectives of metadata management by providing a framework for maintaining the quality and integrity of data assets. It involves defining roles and responsibilities, establishing data standards, and ensuring compliance with regulations, all of which contribute to the overall success of data-driven strategies [13].

In the context of big data, effectively managing metadata involves addressing the quality of the metadata itself. As data is collected in big data settings, the metadata tends to be unstructured, diverse, and may contain inconsistent values. This lack of structure and consistency can hinder traditional data management practices and adversely impact data processing operations. Although the management of metadata and the quality of metadata are closely intertwined, they are often addressed as separate concerns. Consequently, this study places significant emphasis on exploring methods to improve metadata quality at various stages within a big data environment. The paper also delves into the establishment of successful metadata governance and concludes by pointing towards potential avenues for future research in this domain.

Embedding metadata into digital files proves to be an efficient approach for simplifying file management, particularly in the context of driving. Embedded metadata comprises crucial information about the digital asset within the file itself, facilitating the seamless location of images or audio files on your computer. This obviates the necessity to recollect lengthy URLs or grapple with retrieving the original source when using or publishing your images. Furthermore, it ensures that individuals intending to use your images are duly informed of any copyright restrictions and equipped with contact details, thereby simplifying the process of granting permission and upholding ownership rights. Additionally, embedded metadata engenders features that streamline and enhance research and presentations by rendering images, video, and audio files searchable and shareable, with

[12] Amarmeet Kaur et al., "Literature Review on Metadata Governance Article History," *Open International Journal of Informatics (OIJI)* 11, no. 1 (2023): 114–120.
[13] Mohammed Khader and Marcel Karam, "Assessing the Effectiveness of Masking and Encryption in Safeguarding the Identity of Social Media Publishers from Advanced Metadata Analysis," *Data* 8, no. 6 (2023): 1–22.

readily identifiable content, sources, and rights. The strategic integration of embedded metadata serves to enhance the organization, accessibility, and safeguarding of your digital assets while on the move[14].

As the volume of data being generated continues to grow exponentially, the importance of metadata creation has also increased significantly. Metadata serves the crucial role of enabling improved data analytics, data integration, and resource management by providing valuable insights. It also plays a critical role in understanding the context and quality of the underlying data. However, challenges arise due to deviations from metadata recording standards. These may include incomplete attribute information, missing publication URLs, and inadequate provenance documentation. Inconsistent recorded metadata, such as varying value formats, the presence of special characters, and inaccurately inputted values, can further complicate data management. Addressing these inconsistencies through meticulous metadata preparation is vital for enhancing the overall user experience and ensuring the reliability and usability of the data.

Interoperability of Metadata Communications

Interoperability refers to the ability of different systems or components to communicate with each other, exchange data, and utilize the exchanged information. It is a complex concept that can be categorized into various levels or layers, depending on the specific components involved and the depth of interoperability required. The elements involved in interoperability may vary slightly depending on definitions but typically include a distinction between lower-level technical elements and higher-level organizational elements. This understanding is consistent with the technical, syntactic, semantic, and organizational aspects of interoperability, and this section provides a comprehensive summary of these facets.

Interoperability in a technical context refers to the capacity of diverse systems to efficiently exchange data. This process entails the creation of communication channels and protocols to facilitate the seamless transmission of data, for example, transferring files from a USB stick to a computer. While achieving technical interoperability is generally straightforward in today's digital networks and communication protocols, simply moving data from one point to another is not sufficient. In order to fully harness the potential of the data and derive meaningful insights, it is essential to also address syntactic and semantic interoperability. Syntactic interoperability ensures that data formats and structures are compatible between systems, while semantic interoperability ensures that the exchanged data is correctly interpreted and understood by the receiving system[15]. These

[14] Mukesh choudhary, Anshuman v ramani, and vishwas bhardwaj, "The Significance of Metadata and Video Compression for Investigating Video Files on Social Media Forensic," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology* 3307, no. 3 (2023): 304–313.

[15] Stijn Peeters and Sal Hagen, "The 4CAT Capture and Analysis Toolkit: A Modular Tool for Transparent and Traceable Social Media Research," *Computational Communication Research* 4, no. 2 (2022): 571–589.

aspects of interoperability play a vital role in enabling smooth data exchange and effective utilization of information across diverse systems.

The concept of interoperability extends to organizations, laws, and regulations at the top level. The objective of data exchange between health IT systems is to facilitate more efficient collaboration among communication professionals, rather than being an end in itself. This requires standard business procedures and workflows to enable seamless communication provision across establishments. In Islamic higher education, where various stakeholders have diverse interests, policies are needed to incentivize the exchange of interoperable data and, if necessary, to mandate interoperability through legal directives

METHODOLOGY

In our quest to gain a deeper comprehension of the intricate communication patterns observed on various social media platforms, we diligently embarked on the task of constructing an intricate metadata model[16]. By leveraging the methodologies of Design Research (DR), our aim was to cultivate a comprehensive understanding of the multifaceted dynamics in play [17]. The pivotal role played by Action Design Research (ADR) methodology significantly influenced our approach. ADR's emphasis on the iterative development of artifacts and applications spanning diverse usage scenarios ensured the widespread applicability and effectiveness of our findings. By deliberately adopting this approach, we mitigated the potential impact of individual applications, thereby reinforcing the robustness and reliability of our conclusions. Embracing a collaborative mindset, we fostered a deep partnership between academia and industry, leveraging the collective expertise, knowledge, and insights from both arenas [18]. This collective effort facilitated the joint creation of a resilient and tailored solution designed to address the prevalent communication challenges in the landscape of social media [19].

Throughout the first phase of the project, the diverse team of researchers and data analysts dedicated a significant amount of time and effort to conducting a thorough survey of the most current scholarly articles, academic journals, and industry publications relevant to their specific

[16] Ashok Kumar Mohan, Sethumadhavan Madathil, and K. V. Lakshmy, "Holistic Analytics of Digital Artifacts: Unique Metadata Association Model," *International Journal of Digital Crime and Forensics* 13, no. 5 (2021): 78–100.
[17] Beverly Pasian and Rodney Turner, *Design Methods and Practices for Research of Project Management*, *Design Methods and Practices for Research of Project Management* (Routledge, 2024).
[18] Mark E. Phillips, Oksana L. Zavalina, and Hannah Tarver, "Using Metadata Record Graphs to Understand Digital Library Metadata," *Proceedings of the International Conference on Dublin Core and Metadata Applications*, 2019.
[19] Orchida Fayez Ismail, "Constructing a Metadata Model to Capture the Deconstruction of Literary Characters: Remediation of Evil," *Cogent Arts and Humanities* 8, no. 1 (2021): 1–15, https://doi.org/10.1080/23311983.2021.1973649.

research area. This involved carefully examining a wide range of social media platforms and systematically sifting through the vast amount of data available on these platforms [20]. Additionally, the team actively integrated insights obtained from previous research efforts focused on effective data management practices, incorporating these valuable findings into their project framework [21]. A key focus of their efforts centered around identifying specific attributes that would facilitate the smooth integration of a institution-wide inventory system based on standardized product descriptions. Furthermore, they took on the task of developing a meticulously detailed and robust framework for outlining governance processes within the organization [22]. Building on the comprehensive analysis, the team then proceeded to meticulously construct an initial metadata model tailored to social media, drawing directly from their extensive research findings[23].

In the process of identifying the key elements necessary to effectively characterize pattern communication on social media, we conducted a thorough review of the existing literature in this domain. Our focus was on reputable platforms such as Springerlink, DBLP Computer Science Bibliography, Elsevier, and ScienceDirect, from which we gathered pertinent resources [24]. Our selection criteria prioritized works that specifically addressed data protection and data source descriptions, as well as those that delved into the economic aspects of data. It is important to note that our review was confined to English or German literature, and excluded works in other languages. Through our investigation, we identified two significant vocabularies: the Data Catalog Vocabulary (DCAT) and the Asset Description Metadata Schema (ADMS), both of which played a crucial role in achieving our research objectives. While DCAT provides an extensive overview of data sources, ADMS serves as a specialized metamodel designed specifically for describing vocabularies [25].

Figure 1. Workflow of Modeling of Metadata in Social Media

[20] Karen M. Wickett, "Metadata as a Site for Critical Inquiry," in *Proceedings of the International Conference on Dublin Core and Metadata Applications*, 2023, 1–10.
[21] Milagros Maria Erazo-Moreno et al., "Social Media and Education: Perspectives on Digital Inclusion in the University Setting," *Data and Metadata* 3, no. 2 (2024): 1–9.
[22] Nathan Fox et al., "Enriching Social Media Data Allows a More Robust Representation of Cultural Ecosystem Services," *Ecosystem Services* 50, no. September 2020 (2021): 101328, https://doi.org/10.1016/j.ecoser.2021.101328.
[23] Protecting Communication Metadata, *Analyzing and Protecting Communication Metadata*, 2021.
[24] Manuel A. Vázquez et al., "Validation of Scientific Topic Models Using Graph Analysis and Corpus Metadata," *Scientometrics* 127, no. 9 (2022): 5441–5458, https://doi.org/10.1007/s11192-022-04318-5.
[25] Jane Greenberg et al., "Metadata as Data Intelligence," *Data Intelligence* 5, no. 1 (2023): 1–5.

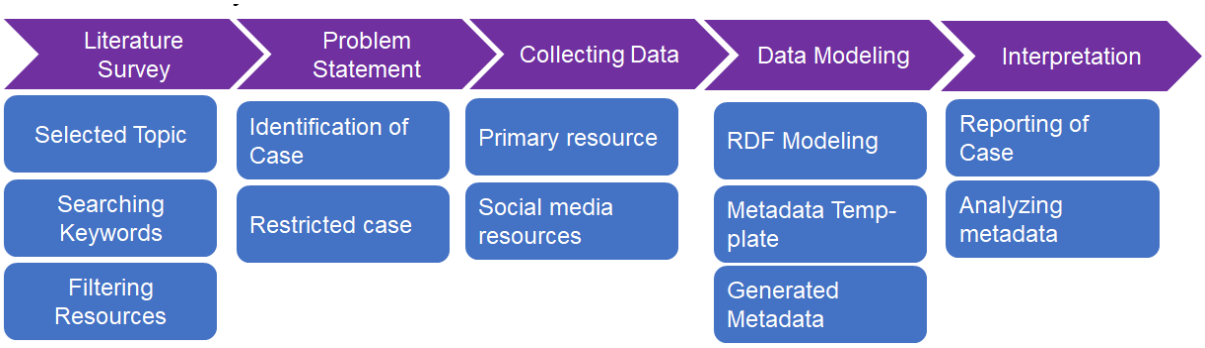


Figure 1. Workflow of Modeling of Metadata in Social Media

The work method outlined is based on the image provided and comprises 5 key stages that must be completed in a cyclical manner. The process commences with a thorough examination of a variety of literature sources that address metadata pertaining to the dissemination of messages on social media. During the initial phase of this review, our focus is on the identification, selection, and elaboration of the problem topic, specifically concentrating on the analysis of communication patterns associated with metadata in message delivery on social media platforms. The choice of the topic concerning message delivery metadata is informed by an extensive search for pertinent keywords to discern the issues surrounding message metadata. We have collated data from 240 pieces of registered literature and have conducted a meticulous and comprehensive selection process[26].

As we progress to the second stage of our process, we meticulously delve into the issue at hand, meticulously examining it from both a broad, overarching perspective, as well as delving into specific details. Our primary focus is to analyze the messaging patterns and content across various social media platforms, giving particular emphasis to platforms like Twitter. Moving into the third stage, we make a strategic decision to concentrate our analysis specifically on Twitter, as we consider it to be our preferred medium for data collection. Our aim is to carry out a thorough analysis that does not rely on visual content, thus streamlining our project workflow. It's critical to highlight that identifying social media messages involves dealing with a vast and diverse range of data that currently lacks definitive metadata standards.

Upon gaining a thorough comprehension of the specific issue at hand, we progress to the data collection phase, which involves engaging with various relevant stakeholders. For instance, we have elected to prioritize Twitter as the primary platform for both media and data collection due to the extensive dataset it offers for our analysis. In addition to its traditional use as a social media platform, Twitter serves as a valuable digital resource

[26] Vanessa Irvin and Bard, "A Whole New Information World: AI, Bots, Metadata, and Discourse," *International Journal of Information, Diversity and Inclusion* 7, no. 1–2 (2023): 1–6.

for the dissemination of written communication, particularly focusing on text-based messaging. This feature greatly facilitates our tracking of interactions, which is essential for our analysis[27].

During the fourth stage of our process, we meticulously engage in a thorough series of modeling activities, meticulously analyzing carefully gathered and delineated data. We have a wide range of tools at our disposal to meticulously identify the most suitable models that adhere to metadata standards. After going through a rigorous selection process, we have opted for a prevalent and appropriate modeling tool, excluding the Dublin Core framework, to facilitate our modeling activities.

One of our primary approaches to modeling data involves strategically utilizing and optimizing the Resource Description Framework (RDF). As we near completion of the RDF modeling process, we are dedicating ourselves to the diligent development and compilation of a metadata template, which will serve as a repository for the simulated data. These meticulously crafted data templates or metadata are purposefully designed to comprehensively meet the intricate information requirements, thereby facilitating the provision of a thorough description of each message delivery entity. Following successful deployment and simulation testing of the metadata template, the resulting metadata will be presented at the conclusion of the process, offering rich, descriptive information about an entity and its attributes[28].

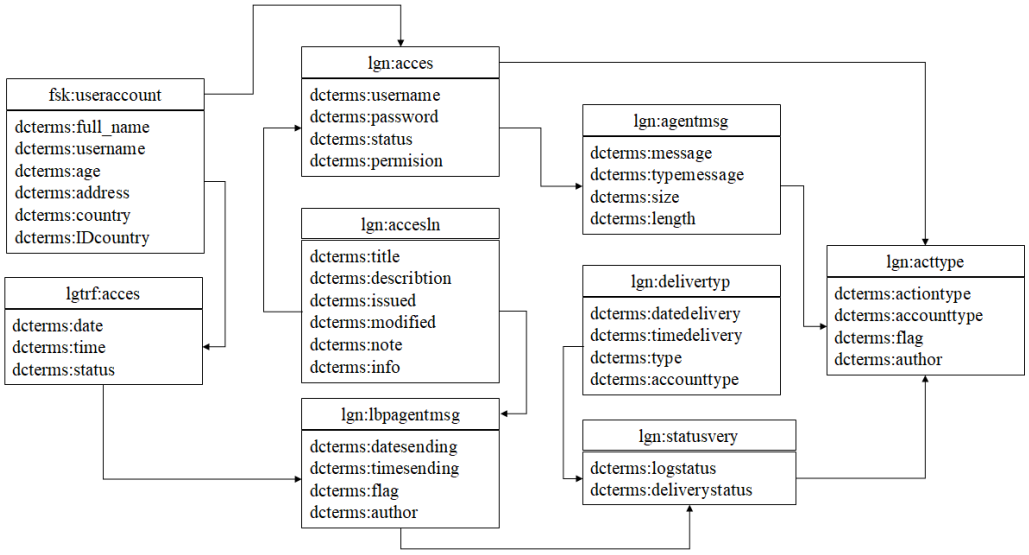
In the culminating stage of the process, our primary focus is on precisely defining and showcasing metadata [29]. This encompasses the visualization of the outcomes of creating and structuring metadata containing data sourced from a variety of social media platforms. We meticulously analyze distinct communication patterns using the Twitter application, both as the medium for communication and as the focal point of our analysis [30]. Following this, we undertake a thorough feasibility analysis to uncover and tackle any potential challenges. For instance, if we were engaged in a metadata production endeavor, we would present our research findings within the framework of interoperability. Furthermore, our findings would be presented in the form of a wide range of real-life case studies and information to offer a comprehensive perspective [31].

RESULT AND DISCUSION

[27] Jonathan Furner, "Definitions of 'Metadata': A Brief Survey of International Standards," *Journal of the Association for Information Science and Technology* 71, no. 6 (2020): E33–E42.
[28] Stacy Allison-Cassin and Dean Seeman, "Metadata as Knowledge," *KULA: Knowledge Creation, Dissemination, and Preservation Studies* 6, no. 3 (2022): 1–4.
[29] Christophe Bahim et al., "The FAIR Data Maturity Model: An Approach to Harmonise FAIR Assessments," *Data Science Journal* 19, no. 1 (2020): 1–7.
[30] Danijela Jaksic, Sanja Candrlic, and Patrizia Poscic, "From User Requirements to Document Repository Enriched with Metadata - A Case Study," in *Procedia Computer Science*, vol. 204 (Elsevier B.V., 2022), 760–767, https://doi.org/10.1016/j.procs.2022.08.092.
[31] Moritz Lehne et al., "Why Digital Medicine Depends on Interoperability," *npj Digital Medicine* 2, no. 1 (2019): 1–5, http://dx.doi.org/10.1038/s41746-019-0158-1.

The outcome is a metadata model that is based on the ADMS and includes the essential characteristics for managing data. This metadata model, as depicted in Figure 1, was created by examining the existing patterns of communication in social media, engaging in discussions with companies, and analyzing existing literature [32]. During the design phase, we added essential characteristics and removed properties that were irrelevant to the specific context. Whenever feasible, we utilized existing vocabularies to precisely define and categorize the model's properties. Through the utilization of the 3-levels repository, asset, and distribution, as well as the elimination and consolidation of individual properties into classes and themes, the data model maintains its extensibility for future versions [33].

Figure 2. Metadata Modeling in Social Media



The primary objective of the metadata model is to provide a standardized description of data sources, treating them as valuable assets within organizations [34]. By conceptualizing data as a product, it becomes essential to manage it as diligently as any other tangible [35]. In collaboration with a prominent global partner in the pharmaceutical industry, extensively involved in procuring third-party data (such as patient records and drug studies), we have designed an inventory software [36]. This software facilitates the internal and external management of data assets and assists in

[32] Thiago Adriano Coleti et al., "TR-Model. A Metadata Profile Application for Personal Data Transparency," *IEEE Access* 8, no. 1 (2020): 75184–75209.
[33] Will Gregg et al., "A Literature Review of Scholarly Communications Metadata," *Research Ideas and Outcomes* 5, no. 2 (2019): 1–13.
[34] Sabina Leonelli et al., "From FAIR Data to Fair Data Use: Methodological Data Fairness in Health-Related Social Media Research," *Big Data and Society* 8, no. 1 (2021): 1–14.
[35] Luca M. Ghiringhelli et al., "Shared Metadata for Data-Centric Materials Science," *Scientific Data* 10, no. 1 (2023): 1–18.
[36] Dominique Batista et al., "Machine Actionable Metadata Models," *Scientific Data* 9, no. 1 (2022): 1–8.

establishing data governance processes, including the assignment of responsibilities and assessment[37]. Our software prototype, developed as a web application, encourages users to compile existing data sources, search for registered metadata, and visualize statistics and indicators related to inventoried data sources[38].

The procedure entails the consolidation of all data sources utilized within a institution and their storage in a centralized repository. Subsequently, these sources are categorized according to a standardized model. This consistent approach in describing and managing data sources offers various benefits. It facilitates the identification and removal of duplicate data through an internal search function, streamlines the identification of alternative data sources based on specific entities, and provides easier control options, including user evaluation. Furthermore, the inventory can be expanded to encompass external data sources from data marketplaces and industrial platforms via a standardized interface.

CONCLUSION

The growing complexity of data handling and management poses significant challenges for businesses. As a response to these challenges, a new metadata model known as the Metadata Model for Pattern Communication in social media has been developed. This innovative model aims to provide a comprehensive framework for describing data sources within an economic context. During the development phase, extensive research was conducted to analyze and compare existing descriptions of metadata on social media. This thorough examination led to the identification of related metadata models, which served as the basis for the standardized design of the newly proposed model. Building upon the principles of the Asset Description Metadata Schema (ADMS), the model was further enriched by incorporating additional properties and attributes, which were then organized using the concept of classes.

In order to ensure practicality and relevance, the model underwent a rigorous evaluation process and was refined based on valuable insights gathered from workshops with a diverse set of companies. Subsequently, the effectiveness of the Metadata Model for Pattern Communication in Social Media was put to the test by integrating it as a database schema for an international institution with substantial data requirements [39]. This validation process unequivocally confirmed the model's efficiency in effectively managing and organizing a diverse range of data sources[40].

---

[37] Derrick P. Snowden et al., "Data Interoperability between Elements of the Global Ocean Observing System," *Frontiers in Marine Science* 6, no. JUL (2019): 1–15.

[38] Larysa Visengeriyeva and Ziawasch Abedjan, "Anatomy of Metadata for Data Curation," *Journal of Data and Information Quality* 12, no. 3 (2020).

[39] Rahul Ramachandran, Kaylin Bugbee, and Kevin Murphy, "From Open Data to Open Science," *Earth and Space Science* 8, no. 5 (2021): 1–17.

[40] Fernando Filgueiras and Lizandro Lui, "Designing Data Governance in Brazil: An Institutional Analysis," *Policy Design and Practice* 6, no. 1 (2023): 41–56, https://doi.org/10.1080/25741292.2022.2065065.

In our upcoming research initiatives, we are highly interested in utilizing the Metadata Model For Pattern Communication In The Social Media as a foundational framework for establishing a standardized description of data sources. We are excited about the opportunity to collaborate with our current partners and form new collaborations with companies in various industry sectors. Our primary goal is to thoroughly evaluate and enhance the design and overall applicability of the model through successive iterations aligned with the Design Science Research (DSR) framework. One potential avenue for progress involves broadening the data foundation by exploring recently established data marketplaces. We believe that previously overlooked data marketplaces, which were dismissed due to limited availability of information, deserve a second look. We understand that the selection of marketplaces is a critical aspect and are dedicated to refining this process in our future research. Through our efforts, we hope to inspire the broader community to pursue further exploration and investigation in this field.

REFERENCES

Alhassun, Atheer S., and Murad A. Rassam. "A Combined Text-Based and Metadata-Based Deep-Learning Framework for the Detection of Spam Accounts on the Social Media Platform Twitter." *Processes* 10, no. 3 (2022).

Ali, Adnan, Jinlong Li, Huanhuan Chen, Uzair Aslam Bhatti, and Asad Khan. "Real-Time Spammers Detection Based on Metadata Features with Machine Learning." *Intelligent Automation and Soft Computing* 38, no. 3 (2023): 241–258.

Allison-Cassin, Stacy, and Dean Seeman. "Metadata as Knowledge." *KULA: Knowledge Creation, Dissemination, and Preservation Studies* 6, no. 3 (2022): 1–4.

Bahim, Christophe, Carlos Casorrán-Amilburu, Makx Dekkers, Edit Herczog, Nicolas Loozen, Konstantinos Repanas, Keith Russell, and Shelley Stall. "The FAIR Data Maturity Model: An Approach to Harmonise FAIR Assessments." *Data Science Journal* 19, no. 1 (2020): 1–7.

Batista, Dominique, Alejandra Gonzalez-Beltran, Susanna Assunta Sansone, and Philippe Rocca-Serra. "Machine Actionable Metadata Models." *Scientific Data* 9, no. 1 (2022): 1–8.

Beretta, Valentina, Jean Christophe Desconnets, Isabelle Mougenot, Muhammad Arslan, Julien Barde, and Véronique Chaffard. "A User-Centric Metadata Model to Foster Sharing and Reuse of Multidisciplinary Datasets in Environmental and Life Sciences." *Computers and Geosciences* 154, no. July 2019 (2021): 1–10.

Coleti, Thiago Adriano, Pedro Luiz Pizzigatti Corrêa, Lucia Vilela Leite Filgueiras, and Marcelo Morandini. "TR-Model. A Metadata Profile Application for Personal Data Transparency." *IEEE Access* 8, no. 1 (2020):

75184–75209.

Erazo-Moreno, Milagros Maria, Gloria María Villa-Córdova, Geraldine Amelia Avila-Sánchez, Fabiola Kruscaya Quispe-Anccasi, Segundo Sigifredo Pérez-Saavedra, and Jhonny Richard Rodriguez-Barboza. "Social Media and Education: Perspectives on Digital Inclusion in the University Setting." *Data and Metadata* 3, no. 2 (2024): 1–9.

Fayez Ismail, Orchida. "Constructing a Metadata Model to Capture the Deconstruction of Literary Characters: Remediation of Evil." *Cogent Arts and Humanities* 8, no. 1 (2021): 1–15. https://doi.org/10.1080/23311983.2021.1973649.

Filgueiras, Fernando, and Lizandro Lui. "Designing Data Governance in Brazil: An Institutional Analysis." *Policy Design and Practice* 6, no. 1 (2023): 41–56. https://doi.org/10.1080/25741292.2022.2065065.

Fox, Nathan, Laura J. Graham, Felix Eigenbrod, James M. Bullock, and Katherine E. Parks. "Enriching Social Media Data Allows a More Robust Representation of Cultural Ecosystem Services." *Ecosystem Services* 50, no. September 2020 (2021): 101328. https://doi.org/10.1016/j.ecoser.2021.101328.

Furner, Jonathan. "Definitions of 'Metadata': A Brief Survey of International Standards." *Journal of the Association for Information Science and Technology* 71, no. 6 (2020): E33–E42.

Ghiringhelli, Luca M., Carsten Baldauf, Tristan Bereau, Sandor Brockhauser, Christian Carbogno, Javad Chamanara, Stefano Cozzini, et al. "Shared Metadata for Data-Centric Materials Science." *Scientific Data* 10, no. 1 (2023): 1–18.

Greenberg, Jane, Ming Fang Wu, Wei Liu, and Fenghong Liu. "Metadata as Data Intelligence." *Data Intelligence* 5, no. 1 (2023): 1–5.

Gregg, Will, Christopher Erdmann, Laura Paglione, Juliane Schneider, and Clare Dean. "A Literature Review of Scholarly Communications Metadata." *Research Ideas and Outcomes* 5, no. 2 (2019): 1–13.

Heacock, Michelle L., Adeline R. Lopez, Sara M. Amolegbe, Danielle J. Carlin, Heather F. Henry, Brittany A. Trottier, Maria L. Velasco, and William A. Suk. "Enhancing Data Integration, Interoperability, and Reuse to Address Complex and Emerging Environmental Health Problems." *Environmental Science and Technology* 56, no. 12 (2022): 7544–7552.

Hippen, Ariel A., and Casey S. Greene. "Expanding and Remixing the Metadata Landscape." *Trends in Cancer* 7, no. 4 (2021): 276–278. https://doi.org/10.1016/j.trecan.2020.10.011.

Irvin, Vanessa, and Bard. "A Whole New Information World: AI, Bots, Metadata, and Discourse." *International Journal of Information, Diversity and Inclusion* 7, no. 1–2 (2023): 1–6.

Jaksic, Danijela, Sanja Candrlic, and Patrizia Poscic. "From User

Requirements to Document Repository Enriched with Metadata - A Case Study." In *Procedia Computer Science*, 204:760–767. Elsevier B.V., 2022. https://doi.org/10.1016/j.procs.2022.08.092.

Jung, Woojin, Donghun Kim, Seojin Nam, and Yongjun Zhu. "Suicidality Detection on Social Media Using Metadata and Text Feature Extraction and Machine Learning." *Archives of Suicide Research* 27, no. 1 (2023): 13–28. https://doi.org/10.1080/13811118.2021.1955783.

Kaur, Amarmeet, Medzhakhed Usama, Nur Liana, Doris Hooi-Ten Wong, and Nurazean Maarop. "Literature Review on Metadata Governance Article History." *Open International Journal of Informatics (OIJI)* 11, no. 1 (2023): 114–120.

Khader, Mohammed, and Marcel Karam. "Assessing the Effectiveness of Masking and Encryption in Safeguarding the Identity of Social Media Publishers from Advanced Metadata Analysis." *Data* 8, no. 6 (2023): 1–22.

Lehne, Moritz, Julian Sass, Andrea Essenwanger, Josef Schepers, and Sylvia Thun. "Why Digital Medicine Depends on Interoperability." *npj Digital Medicine* 2, no. 1 (2019): 1–5. http://dx.doi.org/10.1038/s41746-019-0158-1.

Leonelli, Sabina, Rebecca Lovell, Benedict W. Wheeler, Lora Fleming, and Hywel Williams. "From FAIR Data to Fair Data Use: Methodological Data Fairness in Health-Related Social Media Research." *Big Data and Society* 8, no. 1 (2021): 1–14.

Metadata, Protecting Communication. *Analyzing and Protecting Communication Metadata*, 2021.

Mohan, Ashok Kumar, Sethumadhavan Madathil, and K. V. Lakshmy. "Holistic Analytics of Digital Artifacts: Unique Metadata Association Model." *International Journal of Digital Crime and Forensics* 13, no. 5 (2021): 78–100.

Mukesh choudhary, Anshuman v ramani, and vishwas bhardwaj. "The Significance of Metadata and Video Compression for Investigating Video Files on Social Media Forensic." *International Journal of Scientific Research in Computer Science, Engineering and Information Technology* 3307, no. 3 (2023): 304–313.

Musen, Mark A., Martin J. O'Connor, Erik Schultes, Marcos Martínez-Romero, Josef Hardi, and John Graybeal. "Modeling Community Standards for Metadata as Templates Makes Data FAIR." *Scientific Data* 9, no. 1 (2022): 1–15.

Pacheco, Diogo, Pik-Mai Hui, Christopher Torres-Lugo, Bao Tran Truong, Alessandro Flammini, and Filippo Menczer. "Uncovering Coordinated Networks on Social Media: Methods and Case Studies." *Proceedings of*

the International AAAI Conference on Web and Social Media 15, no. 1 (2021): 455–466.

Pasian, Beverly, and Rodney Turner. *Design Methods and Practices for Research of Project Management*. *Design Methods and Practices for Research of Project Management*. Routledge, 2024.

Peeters, Stijn, and Sal Hagen. "The 4CAT Capture and Analysis Toolkit: A Modular Tool for Transparent and Traceable Social Media Research." *Computational Communication Research* 4, no. 2 (2022): 571–589.

Phillips, Mark E., Oksana L. Zavalina, and Hannah Tarver. "Using Metadata Record Graphs to Understand Digital Library Metadata." *Proceedings of the International Conference on Dublin Core and Metadata Applications*, 2019.

Ramachandran, Rahul, Kaylin Bugbee, and Kevin Murphy. "From Open Data to Open Science." *Earth and Space Science* 8, no. 5 (2021): 1–17.

Rutherford, Brienna N., Carmen C.W. Lim, Benjamin Johnson, Brandon Cheng, Jack Chung, Sandy Huang, Tianze Sun, Janni Leung, Daniel Stjepanović, and Gary C.K. Chan. "#TurntTrending: A Systematic Review of Substance Use Portrayals on Social Media Platforms." *Addiction* 118, no. 2 (2023): 206–217.

Snowden, Derrick P., Vardis Tsontos, Nils Olav Handegard, Marcos Zarate, Kevin M. O'Brien, Kenneth S. Casey, Neville Smith, et al. "Data Interoperability between Elements of the Global Ocean Observing System." *Frontiers in Marine Science* 6, no. JUL (2019): 1–15.

Vázquez, Manuel A., Jorge Pereira-Delgado, Jesús Cid-Sueiro, and Jerónimo Arenas-García. "Validation of Scientific Topic Models Using Graph Analysis and Corpus Metadata." *Scientometrics* 127, no. 9 (2022): 5441–5458. https://doi.org/10.1007/s11192-022-04318-5.

Visengeriyeva, Larysa, and Ziawasch Abedjan. "Anatomy of Metadata for Data Curation." *Journal of Data and Information Quality* 12, no. 3 (2020).

White, Hollie C., Leisa Gibbons, and Eileen Horansky. "Semantic Metadata as Meaning Making: Examining #hashtags and Collection Level Metadata." In *Proceedings of the International Conference on Dublin Core and Metadata Applications*, 2:37–41, 2019.

Wickett, Karen M. "Metadata as a Site for Critical Inquiry." In *Proceedings of the International Conference on Dublin Core and Metadata Applications*, 1–10, 2023.

Zhang, Guilan, Jian Wang, Jianping Liu, and Yao Pan. "Relationship between the Metadata and Relevance Criteria of Scientific Data." *Data Science Journal* 20, no. 1 (2021): 1–21.