



Optimizing Metadata Creation in Institutional Repositories: Best Practices for Enhancing Accessibility, Discoverability, and Reusability

Dr. Ramadan Elaiees

Dept. of Information studies, University of Benghazi, Libya.

ABSTRACT

Metadata is critical to the success of digital repositories since it serves as the foundation for the functioning and usability of institutional repositories. Metadata is used to define digital items so that they may be found, preserved, and managed more easily. Quality metadata is critical for building a successful institutional repository because it allows users to quickly and properly discover and access digital objects of interest. The purpose of this article is to investigate the significance of metadata quality in institutional repositories, as well as the principles and best practices for generating successful metadata. It is commonly established that the quality of metadata in digital repositories influences the repository's success, hence it is critical to guarantee that information is of high quality. Therefore, in order to create a successful repository, it is vital to fully understand the guiding principles and best practices for creating useful metadata. In terms of methodology, various tools have been used and applied to accomplish the objective of this paper such as personal readings of intellectual production, literature review, experience, and the insights of other experts and specialists. The paper concludes that metadata is of great importance for institutional repositories as it is required for effective digital resource management, discovery, and reuse as well as for digital resource interchange, preservation, and measurement of impact.

KEYWORDS: Metadata, Metadata quality, Institutional repositories, Digital repositories, Accessibility, Discoverability, Reusability

INTRODUCTION

There are significant research and development efforts in digital repositories around the world. The quality of metadata is regarded as extremely critical for the primary functions of a digital repository (Tsiflidou, E., Manouselis, N. 2013). Institutional repositories (IR) are digital collections of digital content and metadata maintained by a specific institution, such as a university, research center, or government agency. IRs offer a unique chance to conserve and promote an institution's intellectual output, facilitating information sharing and improving the global awareness of its research and achievements. These collections serve as a repository for institutional knowledge and research output (Schöpfel, J., & Azeroual, O. 2021). Metadata plays an important role in IRs for organizing and managing content as well as giving information about the content for search, access, and reuse. Metadata can also be used to track usage, impact, and other

metrics. Furthermore, metadata can be used to link similar resources such as datasets, articles, and other digital assets (Johnston, L. R. (2020).

According to Gregg, W.(2019) metadata can be used to offer persistent identifiers, which can assist assure the integrity and validity of digital items, as well as an instrument for citing digital objects. Metadata can assist researchers and other stakeholders obtain a better understanding of how digital items are utilized by recording usage, impact, and other indicators. Furthermore, metadata can be utilized to help ensure that digital items are kept in their original form for extended periods of time (Reilly, M., & Thompson, S. 2020). Metadata is just a description or contextual information that adds value to digital materials. It is information about other data that is stored in a database or a digital library or an institutional repository and it is required for effective digital resource management, discovery, and reuse. It is also

required for digital resource interchange, preservation, and measurement of impact (Johnston, L. R. (2020). Metadata contains critical information about the digital content, such as the author, date created, file size, and so on. This allows users to swiftly and readily identify and access the digital resource, as well as track and analyze their usage and impact (Schöpfel, J., & Azeroual, O. (2021).

The quality of metadata is critical for digital libraries, institutional repositories, subject repositories, learning object repositories, and the larger web environment (Schöpfel, J., & Azeroual, O. 2021). Overall, metadata is essential for effective digital resource management, discovery, and reuse. It plays an important role in the organization, management, and preservation of digital resources, and it is essential for measuring the impact of digital objects quickly and easily. It also enables improved interoperability of digital resources across diverse systems and networks

METADATA STANDARDS

There are many metadata standards that have been developed for different purposes, domains, and types of resources. The followings are just some examples of metadata standards. There are many more standards that have been developed for specific communities or applications.

Dublin Core: A simple and general metadata standard that consists of 15 core elements to describe any type of resource. It is widely used for web resources and interoperability across different systems and schemas.

MARC: A metadata standard that was originally designed for cataloging bibliographic records in libraries. It consists of a set of codes and tags to represent various elements and sub elements of bibliographic information.

MODS: A metadata standard that is derived from MARC and provides a more flexible and expressive schema for describing bibliographic resources. It uses XML as its syntax and can accommodate complex and varied resource types.

EAD: A metadata standard that is used for encoding archival finding aids. It provides a hierarchical structure to describe the context, content, and arrangement of archival collections.

DDI: A metadata standard that is used for documenting social science data. It covers aspects such as study design, data collection, data processing, data analysis, data distribution, and data preservation.

ISO 19115: A metadata standard that is used for describing geographic information and services. It covers elements such as spatial and temporal extent, data quality, access and use constraints, and spatial reference system

e-GMS: A metadata standard that is used for describing government information resources. It is based on Dublin Core

and extends it with additional elements and vocabularies to meet the specific needs of the government sector (Guides to Metadata and Discovery, 2023. Available online at: <https://pitt.libguides.com/metadatadiscovery/metadata-standards>)

METADATE IN INSTITUTIONAL REPOSITORIES

Metadata is essential in institutional repositories because it makes digital materials more discoverable and accessible (Schöpfel, J., & Azeroual, O. 2021). It makes it simple for users to access, manage, and organize digital assets. Furthermore, metadata aids in the provision of information about digital assets such as title, author, and publication date. This might be advantageous for users who are looking for specific information or material. Metadata may also be used to characterize digital assets in a variety of ways, such as type, format, subject, or language. This might help users locate the specific material or information they are looking for. As a result, metadata is critical for institutional repositories since it helps to make digital materials more discoverable and accessible (Satija, M. P., Bagchi, M., & Martínez-Ávila, D. 2020). Metadata also improves the discoverability and accessibility of digital objects in institutional repositories besides it makes digital content in institutional repositories easier to discover and access. Accordingly, institutional repositories should not neglect this step. Anne J. Gilliland (2008) stats that “regardless of physical or intellectual form, all information items or objects residing in a digital repository have three characteristics: content, context, and structure, all of which may and should be presented through metadata”.

Content” relates to what the entity comprises or is about and it is vital to an information item.

Context refers to the who, what, why, where, and how variables involved in its production.

Structure: refers to the formal set of relationships inside or between particular information items, which might be internal, external, or both.

Metadata creation and maintenance have grown into a complex blend of human and automated operations and layers created by several activities and persons at various phases of an information object’s life cycle. Metadata management is one emerging area that aims to ensure that the metadata we rely on to validate web resources is trustworthy in and of itself, and that the large volume of metadata that can potentially accumulate throughout the life of a web resource is subject to a summarization and disposition rule. Figure 1 depicts the many stages that information objects often go through over their life cycles in today’s digital environment. As they proceed through their lifecycles, information objects develop layers of metadata that may be linked to them in a variety of ways (Anne J. Gilliland ,2008).

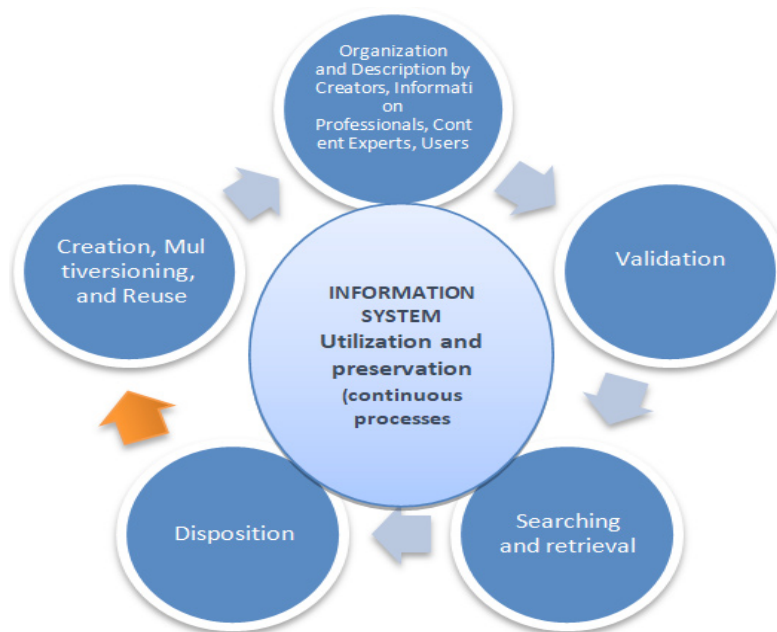


Figure 1. The Life Cycle of an Information Object

Source: Introduction to metadata

METADATE QUALITY

Metadata quality is defined as how accurately metadata describes data and how helpful it is for the intended purposes. Metadata quality can be measured in a variety of ways, including accuracy, completeness, consistency, timeliness, relevance, and accessibility. High-quality metadata assists users in finding, understanding, interpreting, and managing data (Phiri, L. 2020). The practice of analyzing how well metadata fits particular criteria or standards for its intended purposes is known as metadata quality evaluation. Metadata quality can be assessed using a variety of tools and approaches, including validation, analysis, visualization, and feedback. Metadata quality assessments can aid in data discovery, access, use, and management. Schöpfel, J. (2021) states that metadata plays an important role in institutional repositories by providing a standard format for organizing and describing the content on these platforms. Institutional repositories are digital collections of research output produced by institutions, such as universities or research centers. These repositories serve as a means of preserving and disseminating research output, such as articles, data sets, and theses and making it easily accessible and discoverable for academic researchers and the wider public.

Metadata is essentially data that describes other data. In the case of institutional repositories, metadata provides vital information about the research output stored in the repository. This information includes the title of the material, the author(s), abstract, keywords, date of publication, and other relevant information. One of the key benefits of metadata in institutional repositories is that it enables researchers to easily and quickly search for specific research output (Ahammad, N. 2021).

Metadata makes it possible to conduct targeted searches for materials that match specific criteria. For example, a researcher may want to search for research papers in their area of interest that were published in a specific year. With metadata, they can easily filter their search results and find exactly what they are looking for without having to sift through irrelevant materials. Gregg, W. (2019) states that metadata also ensures the accuracy and consistency of the information available in institutional repositories. This is particularly important since repositories can contain a vast amount of research output, and without proper organization and structure, it can be challenging to navigate effectively. Metadata likewise provides a standardized format for organizing and describing the research output, making it possible to maintain a high level of quality and usability (Tsiflidou, E., Manouselis, N. 2013).

COMMON METADATA QUALITY ISSUES IN INSTITUTIONAL REPOSITORIES

The main cause of retrieval challenges in institutional repositories is poor metadata quality. These problems reduce the information's consistency and accuracy in digital repositories. The following are some common metadata quality issues.

Staffing issues: Inadequate personnel or staff may result in inadequate or inconsistent metadata.

Inadequate training: Inadequate training might lead to metadata mistakes and inconsistencies.

Lack of standards: Lack of community standards may result in incompatible metadata among various repositories.

Incomplete metadata: It may be challenging for users to locate and access resources with incomplete metadata.

Poor quality metadata: Poor metadata quality can lead to inefficient content searches, recall of incorrect resources, or no resources at all.

Lack of authority control: The absence of authority control can lead to inconsistencies in names, titles, and subjects

Lack of interoperability: Data sharing across multiple systems may be challenging if the metadata is not interoperable.

METADATA CREATION BEST PRACTICES IN INSTITUTIONAL REPOSITORIES

Many universities and research institutions are spending extensively on the development of institutional repositories, which are becoming increasingly crucial in the academic research scene. Metadata is critical to the success of these repositories, and without it, the usefulness and effectiveness of these platforms can be significantly reduced (Idiegbeyan-Ose, J., I fijeh, G., Iwu-James, J., & Ilogho, J. 2020). The following are some pointers on how to optimize metadata creation in institutional repositories to enhance accessibility, discoverability and reusability.

Consistency: To guarantee consistency and interoperability throughout the repository, controlled vocabularies, defined metadata schemas, and formats should be employed.

Completeness: To characterize digital materials in institutional repositories, descriptive metadata is required. As a result, providing as much information as possible is critical to facilitatediscovery and reuse of the repository's material. Author, title, subject, date, and format information should all be provided.

Accuracy: The repository's content should be appropriately reflected in the metadata. To ensure that metadata is correct, it should be periodically examined and updated.

Accessibility: Users should have easy access to metadata. As a result of this, concise and straightforward instructions on how to access and obtain metadata should be given.

Preservation: Metadata that aids in the preservation of the repository's material should be made available. This might include details about file formats, versioning, and preservation policies.

Interoperability: Metadata should be validated for interoperability with other systems and repositories using standard protocols and metadata schemas such as Dublin core.

Reusability: It is critical that metadata be easily utilized by others. As such, it is vital to use open metadata standards available under open licenses.

User-centered: When establishing and implementing metadata practices, keep the requirements and viewpoints of the repository's users in mind. Solicit user input and feedback to enhance the metadata's usability and usefulness.

Overall, the key to good metadata practices for institutional repositories is to be consistent, complete, accurate, accessible, preservable, interoperable, reusable, and user-centered. By following these best practices, institutions can ensure that their repositories are effective tools for preserving and sharing scholarly and research materials.

CONCLUDING THOUGHTS

Metadata quality is essential for the effective management and discovery of digital resources in institutional repositories. Metadata quality depends on various factors, such as metadata semantics, application consistency, accuracy, completeness, and interoperability. To ensure and improve metadata quality, institutional repositories need to adopt and implement various quality control mechanisms, such as staff training, manual review, metadata guidelines, and metadata generation tools. These mechanisms can help address common metadata problems, such as inconsistent or inaccurate author names, affiliations, or identifiers. By enhancing the quality of metadata in institutional repositories, the visibility and impact of scholarly outputs can be increased.

In summary, metadata is key to the success of institutional repositories. It helps to standardize and organize the research output, allowing researchers to quickly and easily find the materials they need. It also ensures the accuracy and consistency of the information available, ultimately enhancing the usability and effectiveness of institutional repositories. As such, it is essential that institutions invest in the development of robust metadata systems to ensure the long-term success of their repositories.

REFERENCES

1. Ahammad, N. (2021). Quality control (QC) of an institutional repository: a hands-on. *Collection and Curation*, 40(4), 145-152.
2. Amorim, R. C., Castro, J. A., da Silva, J. R., & Ribeiro, C. (2015). A comparative study of platforms for research data management: interoperability, metadata capabilities and integration potential. In *New Contributions in Information Systems and Technologies: Volume 1* (pp. 101-111). Springer International Publishing.
3. Anne J. Gilliland, 2008. Setting the stage. In, *Introduction to metadata*. Edited by Murtha Baca. — 2nd ed. p. cm. Published by the Getty Research Institute, Los Angeles Getty Publications.
4. Fujita, M. S. L., Agustín-Lacruz, C., Tolare, J. B., Terra, A. L., & Bueno-de-la-Fuente, G. (2022). Institutional repositories and knowledge organization: A bibliographic study from Library and Information Science. *Education for Information*, (Preprint), 1-16.
5. Gregg, W., Erdmann, C., Paglione, L., Schneider, J., & Dean, C. (2019). A literature review of scholarly communications metadata. *Research Ideas and Outcomes*, 5, e38698.

6. Guides to Metadata and Discovery @ Pitt: Metadata Standards. Available at: <https://pitt.libguides.com/metadatadiscovery/metadata-standards>
7. Idiegbeyan-Ose, J., Ifijeh, G., Iwu-James, J., & Ilogho, J. (2020). Management of institutional repositories (IR) in developing countries. In *Digital Libraries and Institutional Repositories: Breakthroughs in Research and Practice* (pp. 356-382). IGI Global.
8. Johnston, L. R. (2020). Data discovery and the role of academic institutional data repositories.
9. Phiri, L. (2020). Automatic classification of digital objects for improved metadata quality of electronic theses and dissertations in institutional repositories. *International Journal of Metadata, Semantics and Ontologies*, 14(3), 234-248.
10. Schöpfel, J., & Azeroual, O. (2021). Metadata quality: implications for library and information science professionals. *Library Review* Vol. 54 No. 5, 2005 pp. 295-300. DOI 10.1108/00242530510600543
11. Reilly, M., & Thompson, S. (2020). Understanding Data Management Planning and Sharing: Perspectives for the Social Scientist. *Anthropological Data in the Digital Age: New Possibilities–New Challenges*, 13-30.
12. Satija, M. P., Bagchi, M., & Martínez-Ávila, D. (2020). Metadata management and application. *Library Herald*, 58(4), 84-107.
13. Schöpfel, J., & Azeroual, O. (2021). Current research information systems and institutional repositories: From data ingestion to convergence and merger. In *Future directions in digital information* (pp. 29-37). Chandos Publishing.
14. Tsiflidou, E., Manouselis, N. (2013). Tools and Techniques for Assessing Metadata Quality. In: Garoufallou, E., Greenberg, J. (eds) *Metadata and Semantics Research. MTSR 2013. Communications in Computer and Information Science*, vol 390. Springer, Cham. https://doi.org/10.1007/978-3-319-03437-9_11

Citation: Dr. Ramadan Elaies, "Optimizing Metadata Creation in Institutional Repositories: Best Practices for Enhancing Accessibility, Discoverability, and Reusability", *American Research Journal of Computer Science and Information Technology*, Vol 6, no. 1, 2022, pp. 1-5.

Copyright © 2023 Dr. Ramadan Elaies, This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.