International Journal of Research
Available at https://journals.pen2print.org/index.php/ijr/
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 06 Issue 09
August 2019

Role of Metadata Management

Mar Lar Htun[1], Win Myat Thuzar[2], Thae Thae Han[2], Thida San[3]
[1]Faculty of Computer Science, University of Computer Studies (Pyay), Myanmar
[2,3]Information Technology Supporting and Maintenance Department, University of Computer Studies (Meiktila), Myanmar
marlarhtun19@gmail.com

Abstract:

The more people that enter website and stay there, the greater in the impression Google has that this site is relevant, popular and authoritative for a particular set of keywords or search phrases. And that in turn influences rankings. It all starts with metadata. When the users complain that search doesn't provide relevant results - it's usually a metadata issue. When the users don't know where to look for business-critical information - it's usually an information architecture issue, including metadata. Metadata provides additional information about the content, helping users to find and use it easier. With good quality metadata, not only the usability but also findability of the content will be skyrockets. Metadata is just as important as content.

Keywords

Metadata, Dublin core.

1. Introduction

Users of data must know what the data is before it becomes useful as information. Information about data is referred to as metadata. The simplest definition of metadata is "data about data." But, to be a bit more precise, metadata describes data, providing information like data type, length, textual description, and other characteristics of the data. **Metadata** characterizes data. It is used to provide documentation such that data can be understood and more readily consumed by your organization. Metadata answers the ***who***, ***what***, ***when***, ***where***, ***why***, and ***how*** questions for users of the data. Metadata is an important way to protect resources and their future accessibility. It's a critical concern given the fragility of digital information and its susceptibility to corruption or alteration. For archiving and preservation purposes, it takes metadata elements that track the object's lineage, and describe its physical characteristics and behavior so it can be replicated on technologies in the future.

2. Metadata

Metadata is data about data. Metadata is structured information that explains, describes, or locates the original primary data, or that otherwise makes using the original primary data more efficient. A wide variety of industries use metadata, but for the purposes of digital imaging, there are currently only a few technical structures or schema that are being employed. A schema is a set of properties and their defined meanings, such as the type of value (date, size, URL, or any useful designation).

2.1. Types of Metadata

Metadata falls into three main categories:

- **Descriptive** - used for discovery and identification and including such information as title, author, abstract and keywords.
- **Structural** - shows how information is put together page order to chapters.
- **Administrative** - enables better resource management by showing such information as when and how the resource was created. Two types of administrative metadata are those that deal with intellectual property rights and preservation metadata, used to archive and preserve a resource.

2.2. Management of Metadata

If metadata helps us to understand data, metadata management enables us to *use* the metadata. Metadata creation is time-consuming and expensive. To be truly useful, once stored, metadata must be centrally available and easy to maintain. The primary goals of metadata management are:

- To promote metadata conformity to enable sharing of metadata by an organization's applications. Metadata that is defined for one application can be copied and easily adapted for use by another application.
- To provide a common, centralized method of searching and managing distinct collections of metadata.

Both goals lower the costs of metadata development and maintenance by promoting standardization and reducing redundancy. Furthermore, when these goals are achieved, metadata can provide meaningful and valuable information, for example,

- Impact analysis of technical changes within an organization.
- Comprehensive technical reporting about the organization's application systems.

Impact analysis gauges the effect of a single technical change on all of the applications in an organization. For example, if an organization stores metadata about its computer systems, it can use that metadata to easily determine which applications will be affected by taking a specific server offline. Or, if all applications store client address information in an address object and a change is needed in the way this information is stored – for example, to surface street address, city/state, and country as three separate fields instead of one – the change is easy to identify, make, and propagate.

As electronic data transfers and e-commerce increase, the soundness of the metadata supporting these transactions will become as important as the data itself. Support for industry metadata models and data interchange standards enables organizations to respond quickly and economically to rapidly evolving, external reporting obligations.

2.3. Usage of Metadata

Metadata helps users to find relevant information and discover resources. Metadata also helps to organize electronic resources, provide digital identification, and support the archiving and preservation of resources.

Metadata is created and collected because it enables and improves use of that data. Good metadata can make up for human fallibilities. Metadata also makes text documents easier to find because it explains exactly what the document is about.

Metadata is described by AIIM as the structured information about a document, data or other information content. Metadata is traditionally found in card catalogs in libraries. With the age of digital documents, metadata is used to describe the content of these documents.

2.4. Benefits of Using Metadata

Metadata provides a number of very important benefits to the enterprise, including:

- **Consistency of definitions** Metadata contains information about data that helps reconcile the difference in terminology such as "clients" and "customers," "revenue" and "sales," etc.

- **Clarity of relationships** Metadata helps resolve ambiguity and inconsistencies when determining the associations between entities stored throughout data environment. For example, if a customer declares a "beneficiary" in one application, and this beneficiary is called a "participant" in another application, metadata definitions would help clarify the situation.

- **Clarity of data lineage** Metadata contains information about the origins of a particular data setand can be granular enough to define information at the attribute level; metadata may maintain allowed values for a data attribute, its proper format, location, owner, and steward. Operationally, metadata may maintain auditable information about users, applications, and processes that create, delete, or change data, the exact timestamp of the change, and the authorization that was used to perform these actions.

2.5. Classification of Metadata

Metadata is used for several purposes;

a. describing data for the purposes of data exchange;
b. describing data for the purposes of global access from query (including update) to optimize recall and relevance;
c. describing data for the purposes of query optimization;
d. describing data for the purposes of answer integration and explanation;
e. describing data for the purposes of correct analytical processing or interpretation, representation or visualization.
f. describing the data to overcome multilingualism and multimedia heterogeneities

All of these purposes require that the data be described:

1. such that the resource is constrained formally to ensure integrity;
2. such that the resource is reachable by automated means;
3. such that there is sufficient description for the purposes to utilize the resource.

3. Dublin Core

Dublin Core is an initiative to create a digital "library card catalog" for the Web. Dublin Core is made up of 15 metadata (data that describes data) elements that offer expanded cataloging information and improved document indexing for search engine programs.

The 15 metadata elements used by Dublin Core are: title (the name given the resource), creator (the person or organization responsible for the content), subject (the topic covered), description (a textual outline of the content), publisher (those responsible for making the resource available), contributor (those who added to the content), date (when the resource was made available), type (a category for the content), format (how the resource is presented), identifier (numerical identifier for the content such as a URL), source (where the content originally derived from), language (in what language the content is written), relation (how the content relates to other resources, for instance, if it is a chapter in a book), coverage (where the resource is physically located), and rights (a link to a copyright notice).

Two forms of Dublin Core exist: Simple Dublin Core and Qualified Dublin Core. Simple Dublin Core expresses elements as attribute-value pairs using just the 15 metadata elements from the Dublin Core Metadata Element Set. Qualified Dublin Core increases the specificity of metadata by adding information about encoding schemes, enumerated lists of values, or other processing clues. While enabling searches to be more specific, qualifiers are also more complex and can pose challenges to interoperability.

Each method of recording or transferring Dublin Core metadata has its plusses and minuses. HTML, XML, RDF, and relational databases are among the more common methods.

The Dublin Core Metadata Initiative began in 1995, taking its name from the location of the original workshop, Dublin, Ohio. It has since become international in scope and has representatives from more than 20 countries now contributing. Dublin Core has always held that resource discovery should be independent from the medium of the resource. So, while Dublin Core targets electronic resources, it aims to be flexible enough to help in searches for more traditional formats of data too. Web sites, though, are the most common users of Dublin Core.

4. Related Work

In [2], author is informed by four years of research and 57 in-depth interview data analysis with practicing librarians, researchers, metadata consultants and library users using a constructivist grounded theory method (Alemu, 2014). From the research, four overarching metadata principles, namely, metadata enriching, linking, openness and filtering emerged. The integration of these principles resulted in the emergence of a new theory of digital library metadata; The Theory of Metadata Enriching and Filtering. Author [3] focuses on an entity–attribute–value (EAV) modeling technique to facilitate database interoperation. This technique, historically based on the 'association list' data structure (Winston, 1984), is designed to represent highly heterogeneous data in a simple and uniform fashion. EAV techniques have traditionally been used to represent the extensional part of a database (i.e. data). In [1], the UNT libraries reviewed several metadata initiatives to build an element set appropriate for its digital collections while monitoring the RLG/OCLC efforts toward building a standard metadata element set. In addressing the issues of identifying specific metadata requirements, UNT Libraries attempted to assess the specific characteristics of the existing digital resources. Based on the thorough assessment of the available digital resources, attempts have been made to review current best practices and standards to represent a range of relevant fields. The review pays particular attention to the preservation and management metadata sets, which are needed to support various preservation approaches including migration and emulation. The work at NLA developed a practical model for dealing with the immediate threat of disappearing digital objects, and established a workable distributed archive. Similarly, a number of projects and researches - such as OAIS (Open Archival Information System), CEDARS (CURL Exemplars in Digital Archives), NEDLIB (Networked European Deposit Library), and others - have investigated options for dealing with long-term preservation challenges. Based on the preliminary survey of the existing digital collection and a detailed review of current best practices, we chose to base our recommendation of preservation metadata on a synthesis of various preservation metadata until the OCLC/RLG (2001) completes a national standard.

5. Conclusion

The ability to apply keywords by using extensive metadata currently exists. The metadata lifecycle extends the lifecycle of datasets. Metadata provides information on data and resources. The quality of the metadata directly affects the discoverability and reuse of the resources. Metadata is essential for understanding the structure of information, its quality and its relevance. Metadata is essential in explaining answers from ever more complex information systems. Metadata assists in distilling knowledge from information and data. Metadata assists in multilingualism and in multimedia representations.

6. References

[1] Daniel Gelaw Alemneh," A Metadata Approach to Preservation of Digital Resources: The university of North Texas Libraries Experience", *University of South Carolina, First Monday, Volume 7, Number 8-5August 2002*

[2] Dr Getaneh Alemu," A Theory of Metadata Enriching and Filtering: Challenges and Opportunities to

e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 06 Issue 09
August 2019

Implementation ", *Qualitative and Quantitative Methods in Libraries (QQML) 5: 311- 334, 2016*

[3] Kei Hoi Cheung, "A Metadata Approach to Query Interoperation between Molecular Biology Database", Oxford University Pass, Vol. 14, No. 6 1988, Page 486-497

[4] Tanmay Mondal, "Metadata Creation Methods: A Study", Journal of Advances in Library and Information Science, ISSN: 2277-2219 Vol. 7. No.1. 2018. pp.177-182