

METADATA AND METADATA STANDARD WITH SPECIAL REFERENCE TO DUBLIN CORE

Masiur Rahaman

¹*Research Scholar, Librarian Rani Dhanya Kumari College, Murshidabad, West Bengal India*

Received: 05 Nov 2021

Accepted: 12 Nov 2021

Published: 25 Nov 2021

ABSTRACT

Expanding use of Internet has intensified awareness amongst the information community of the requirement to arrange user-friendly searching and navigation tools that lead to quality information. To provide an index of available items to save users time and network is a crucial part of securing valuable access to internet resources. The vital link between the information creator and the information user is provided by metadata. Metadata means delicately different thing surrounded by the various regulations that use the term. The term which also become a stylish and is often overused. Nevertheless, in the field of library world and considering to information about resources it is extremely useful and description of digital resources. Even in the 21st century, Cutter's 'objects' are still appropriate to all kinds of information packages, beyond just books in the libraries and to several kinds of data standards (i.e, metadata schema, cataloguing standard, etc). Metadata illustrates objects of multimedia content and varieties. The primary function of metadata is helping retrieval of the objects that it illustrates. This paper discusses about metadata, types of metadata, metadata schemas, metadata standards, Dublin Core, Dublin Core elements and their examples.

KEYWORDS: *Metadata, Metadata Schemas, Metadata Standards, Dublin Core, Dublin Core Elements*

1. INTRODUCTION

Metadata which is structured data and states the disciplines of a resource. It is used as a means of describing any object. It has same characteristics as cataloging in the library. The word 'meta' is originated from the Greek word indicating a nature of greater order. Metadata element is formed with several pre-defined elements, denoting resources. Each element can have one or more value. The fundamental link between the information maker and the information users is provided by metadata. Resource discovery is the basic function of metadata. In the world of information, metadata which may comprise of an approved set of data elements with approved semantics, approved syntax and approved rules for communicating the content of the elements. Metadata schema can be classified as bibliographic metadata schema and non-bibliographic metadata schema. The first one is further classified as textual and multimedia media related metadata schema. The Dublin Core aimed at the description of documents.

2. DEFINITIONS OF METADATA

Metadata is occasionally defined literally as “data about data”. It implies an information package e.g., a book, a webpage, etc is data. So, content of the book would be data about data. Metadata is well-defined as ‘information about data’. Although the term is generally explained to mean controlled data about resources that can be applied to help support wide range of processes. These might incorporate for instance, resource explanation and discovery, the management of information resources and their long-term preservation.

Metadata is a progressively universal term that is recognized in different ways by the various professional communication that shape, create, describe, safeguard, and use information techniques and resources. While these groups, and storehouses, information, and communication technologies-come altogether to avail the information age into a reality, it is important that we comprehend the special role of metadata that can play in the growth of operative, commanding, interoperable, ascendable, and conservable cultural inheritance information and record keeping systems. Library metadata comprises indexes, abstract, and catalogue records created conferring to cataloguing guidelines and organizational and content standards such as MARC (Machine Readable Cataloguing), additionally authority forms such as LCSH (Library of Congress Subject Headings) or the AAT (Art and Architecture Thesaurus). Such bibliographic metadata has been created co-operatively and made obtainable since the 1960s to depots and user through computerized systems such as bibliographic utilities, On-line Public Access Catalogue (OPAC), and commercial online databases.

3. TYPES OF METADATA

Library metadata includes catalogue records, indexes, and abstracts. There are three types of metadata for digital object repositories and online resource discovery and maintenance.

3.1 Descriptive Metadata

Metadata used to designate or to identify characteristics of an information package along with analyzing its intellectual contents. The following types of information are there in this category.

- Intellectual organization data – authority control, names, subjects, etc.
- Data identifying an information package –title, author, date of publication, etc.
- Intellectual data access – subject headings, categorization, classification.

3.2 Structural Metadata

Structural metadata refers the ‘the file structure, dataset, or other package of information that is being described. Structural metadata is also highlighted to as technical metadata, display metadata, or use metadata. It provides the following kinds of information.

- Digitization information- scaling ratios, compression ratios.
- Hardware and software documentation.
- Version control – alternate digital format such as HTML on PDF for text, JPG for image.
- Authentication and security data – password methods

Data relating to creation of digital image- date of scan, resolution.

3.3 Administrative Metadata

Metadata used in handling and controlling information resources, formed for decision making, record keeping, and management purpose. It helps monitoring, reproducing, backing up digital information packages. The following information are including

- Rights and reproduction tracking
- Documentation of legal access requirements
- Acquisition information
- Location information
- Selection criteria for digitization
- Information preservation – integrity information, documentation of physical condition.

Administrative Metadata can be categorized three sub-types. They are:

- Preservation metadata – it is required to enable the long –term storage and serviceability of digital content. Preservation metadata elements must consist of structural type, file description, size, software and hardware environment, transformation history, context information.
- Rights and Access Metadata- the information related to access of information packages, who may use them, and for their purposes. Which may includes access categories, copyright statements, usage information and payments options.
- Meta-Metadata- it is data about data about data. It tracks information about the metadata. (Prasad, ARD.2009)

4. METADATA SCHEMAS

Metadata schemas may be described as sets of elements considered to meet the requirements of communities. Most of the schemas are developed for specific types of information.

Each metadata schema should have three basic characteristics. They are – 1) a limited number of elements, 2) the name of each element, and 3) the meaning of each element. There are seven types of metadata schemas. These schemas are discussed below:

4.1 Learning Objects Description and Discovery Schema

With the advancement of online education, learning technology standards are critical to the success of these industries. Metadata facilitates searching, permits learners, authors, and others to explore, recover and assemble reusable learning matters.

Some famous metadata standards for online learning are:

- Education Network Australia (EDNA)
- IEEE learning Object Metadata
- The gateway to Educational Material (GEM)
- Can Core Learning Resource Metadata Initiative.

4.2 Museum and Artwork Metadata Schema

Museum is a rich collection of visual resources. The art works and artifacts give us information about lost civilizations. So, there is a great need to digitize and preserve the resource. Metadata standards are being developed to digitize the art information. Some metadata standards for museum information are listed below:

- RLG REACH Element Set
- Computer Interchange of Museum Information (CIMI)
- Categories for the Description of Works of Art (CDWA)

4.3 Metadata Standards for Thesis and Dissertations

Thesis and dissertations are very important source of information to the research community. To access this information standardized retrieval techniques are needed. Some famous metadata standards for the Electronic Dissertations and Thesis are as follows:

- XMetaDiss
- NDLTD ETD-MS
- E-theses Metadata Set

4.4 Preservation and Archival Metadata Initiatives

Conservation of information in digital method is one of the biggest tasks for the information expert in the 21st century. The rapid growth in the use of computers and internet has made it essential for us to come up with full proof preservation techniques to facilitate long term preservation of digital information. Metadata can take vital role for long term preservation of digital objects. Some well-known digital preservation metadata standards are as follows:

- Data documentation Initiative (DDI)
- Victoria Electronic Records Strategy (VERS) Metadata Schema
- Preservation Metadata: Implementation Strategies (PREMIS)

4.5 Government Metadata Initiatives

Information regarding Government is an important resource for general people, libraries and business institutions. Some standards need to preserve and digitize the government information. Government agencies usually use the following metadata standards:

- Government Information Locator Service (GLIS)
- Australian Government Locator Service (AGLS)

4.6 Multimedia Object description Metadata Schema

Multimedia objects are still pictures, graphics, music, audio, video, multimedia collections, etc. There is no available established resource discovery standard to describe the structure and relationships between the multimedia objects. So, it is very difficult for searching, browsing and archiving those objects. To cope up with this problem the following multimedia metadata standards have been developed:

- MPEG-21 Multimedia Framework
- Multimedia Content Description Interface (MPEG-7)
- Public Broadcast Metadata Dictionary (Taylor, Arlene G.2005)

4.7 Resource Discovery/Description Metadata Schema

This schema provides the same function as good as cataloging. It facilitates the classification of Internet resources using various element sets for structuring the information about a source in a way that describes, explains, locate, or otherwise makes it easier to retrieve the resources. Some resource discovery metadata standards are stated here:

4.7.1 Metadata Encoding and Transmission Standard (METS)

It is a standard within a digital library for programming, descriptive, administrative and organizational metadata regarding objects.

4.7.2 Metadata Object Description Schema (MODS)

The LC has been developed MODS in 2002. It is an XML Schema. Second version MODS was available in 2003. It is a hybrid metadata schema that incorporates both encoding rules and a set of elements. It has 20 top level elements and 47 sub elements. (www.loc.gov/standards/mods/)

4.7.3 Metadata Authority Description Schema (MADS)

It is also an element that may be used to provide metadata about agents, events, and terms like XML schema for an authority. It is a MARC 21 compatible XML format for the type of data carried in records in the MARC authority's format. (www.loc.gov/standards/mads/)

4.7.4 Encoded Archival Description (EAD)

It is a scheme for encoding archival and library finding aid. It allows the users to find out about collections in distant places. The EAD version 2002 is an SGML format.

4.7.5 Text Encoding Initiative (TEI)

It is a global and interdisciplinary standard that enables libraries, museums, publishers, and individual scholars for online research, teaching, and preservation. The TEI encoding scheme consists of multiple modules or DTD fragments known as tag sets. TEI was established in 1987. It has been used extensively in academic libraries for creating digital texts.

4.7.6 Online Information Exchange (ONIX)

It is a standard arrangement that publishers can use to distribute electronic information about their books and other publications. ONIX covers both content of elements and encoding.

4.7.7 Dublin Core (DC)

The development of Dublin Core as a common metadata standard for exercise by libraries, archives, government, and other publishers of on-line information which is a common set of data elements intended to support the search and discovery of an extensive range of information resources over disseminated networks like the web. The Dublin Core is named after its first workshop, held in 1995 in Dublin, Ohio. It has been approved as NISO Standard Z39.85-2001, and ISO approval is sought. The 'Dublin Core Metadata Initiative' (DCMI) is devoted to nurturing the extensive acceptance of interoperable metadata standards and sponsoring the development of dedicated metadata vocabularies.

Normally, this metadata is being used to increase existing means for searching and indexing web-based metadata. DC has been applied with HTML for a number of years. Recently it has been carried out using XML. The Dublin Core Metadata Element Set (DCMES) delivers a semantic vocabulary for highlighting the core information properties, such as description, creator, and date.

4.7.7.1 Dublin Core Elements

The Dublin Core metadata elements can be categorized into three groups. These are –

1. Content and resource Elements — Title, Subject, Description, Source, Language, Relation, and Coverage
2. Intellectual property Elements – Author or Creator, Publisher, Contributor, and Rights
3. Physical manifestation Elements – Date, Type, Format, and Identifier

Audience and Rights Holder are the latest additions in these core elements.

A short description of each element and examples are given below;

Title – Title is the name given by the creator or publisher to the information package. Examples <meta name = "DC. Title" content = Prolegomena to library classification">

<meta name = "DC. Title" content = Library Herald, Vol.2, No. 3">

Subject – Subject is the topic of resource. It is expressed through controlled vocabularies and formal classification scheme. Examples: <meta name = "DC. Subject" content = "Library science"> <meta name = "DC. Subject" scheme = "MESH"

content = "Carcenoma">

Description – It is the textual description of the content of the resource. Examples:

<meta name = "DC. Description" lang = "en" content = The author presents a tutorial introduction to Perl programming examples with extensive examples on regular expressions. He also deals with scripting in Perl">

Source – The work, either print or electronic, from which this object is derived. Examples: <meta name = "DC. Source" content = William Shakespeare's Hamlet">

<meta name = “DC. Source” content = <http://wbcsc.ac.in>>

Language – Language means the language of the intellectual content of the resource. Examples: <meta name = “DC. Language” content = “en”> <meta name = “DC. Language” scheme = “rfc3066” content = “en-US”>

Relation – Relationship to other resources, e.g., chapters in a book, items in a collection, etc. <meta name = “DC. Relation. Is Part of” content = “<http://foo.bar.org/abc/proceedings/2009/>> <meta name = “DC. Relation. References” content = “urn:isbn:978-1-56592-149-6”>

Coverage- Coverage is the spatial locations and temporal duration characteristics of the resource. It is an identification of spatial location, temporal period or justification. Examples: <meta name = “DC. Coverage” content = “US Civil war era; 1861-1865”>

<meta name = “DC. Coverage. Jurisdiction” content = “Commonwealth of Australia”>

Creator – The person, who is responsible for creating the intellectual part of the resource, is called creator. E.g., authors in case of written documents, artists in

case of visual resources. Examples: <meta name = “DC. Creator” content = “Tagore, Rabindranath”>

Publisher – The name of the entity which is responsible for making the resource. E.g.,

Department of an University, University, Publishing House, etc. Examples: <meta name = “DC. Publisher” content = “McGraw-Hill”>

Contributor – A person or organization, who has made significant intellectual contributions to the resource rather than creator, is called contributor.

E.g., editor, illustrator, etc. Examples: <meta name = “DC. Contributor. Editor” content = “Bhattacharyya, G”>

Rights- It is a statement, link or identifier that gives information about rights held in and over the resource. Examples: <meta name = “DC. Rights”

content = <http://foo.far.org/cgi-bin/terms>>

Date – A date is associated with the creation on availability of the resource. The date format should be YYYY-MM-DD. Examples: <meta name = “DC. Date” content = “2010”> <meta name = “DC. Date” content = “2010-11-26”>

Type – The type is the category of resource, like novel, dictionary, poetry, etc. Examples: <meta name = “DC. Type” content = “poetry”>

Format – It is the data format of the resource, e.g., HTML, jpeg, etc. Examples: <meta name = “DC. Format” content = “text/html”>

Identifier – It is the identification number or string of a resource. E.g., URLs, DOI, ISBN, etc. Examples: <meta name = “DC. Identifier” scheme = “ISBN”

content = “978-81-775-8260-4”>

Audience – It is a class of entity for whom the resource is intended. The creator or the publisher or the third party may determine the target audience.

Examples: <meta name = “DC. Audience” content = “undergraduate students”>

Rights holder – A person or organization who has the rights over the resource. Examples: <meta name = “DC. Rights Holder” content = “Chatterjee Brothers Ltd”>

(www.dublincore.org)

5. CONCLUSION

At primary stage when metadata standards were developed, emphasized on the descriptive elements required for discovery, identification, and retrieval. For improvement of retrieval of the information resources, the requirement of metadata has a potential role. When metadata initiatives were developed, the administrative metadata was further focused, especially in the area of rights and preservation. The area of technical metadata in metadata schemas which still does not significantly get consideration. Metadata work is an ongoing process of development organizations. A subcommittee of Technical Committee 46 is lecturing the development of metadata for the applications bibliographic control in the International Organization for Standardization (ISO). Other than the formal standards community organizations several metadata stipulations have been developed are looking for their specifications turned into international standards. This approach is an example of the Dublin Core. In the accessibility of information, the World Wide Web (WWW) has fashioned a revolt. The applications for metadata modern technologies, standards, and best performs are constantly progressing. In our opinion the modest amount of well-defined construction is necessary if we want to realize the purpose of the standard to improve information retrieval.

REFERENCES

1. “Dublin Core”. <http://www.dublincore.org>
2. “Metadata object description schema”. <http://www.loc.gov/standards/mods/>
3. “Metadata transmission and encoding standards”. <http://www.loc.gov/standards/mets/>
4. Prasad, A R D and Devika P Madalli (2009).” *Metadata standards for textual and multimedia content*”, paper presented at the conference of *Open Access to textual and Multimedia Content: Bridging the digital divide*, January, 29-30.
5. Taylor, Arlene G. (2005). *The organization of information*. 2nd ed. London: Libraries Unlimited.
6. <http://www.getty.edu/research/institute/standards/intrometadata/2articles/index.html>.
7. <http://www.library.berkeley.edu/MOA2>
8. <http://www.purl.oclc.org/metadata/Dublin Core>.