

MODULE:- 4 METADATA AND DUBLIN CORE

4.2 INTRODUCTION: Metadata can be defined as a structured data which describes the attribute of the Dublin Core butes and contents of an original document to provide instant access. It is a description of objects such as images, e-texts, multimedia and other electronic files or data about data or information about information. Simple examples of metadata are index, schema, table and view. The other form of metadata are MARC-21 and rules sets used with it such as AACR2 are metadata standards and Traditional library cataloguing.

4.1 Objectives:

After reading this module in detail, it is expected that you will have a clear cut idea:

- The concept and meaning of Metadata
- The purpose and functions of Metadata have been clearly explained citing different functions with suitable examples.
- Further Types of metadata and benefits also stated in an understandable manner.
- In the Introduction, you can understand the reasons for the development of Dublin Core.
- The Initiatives and efforts made by different organizations for the genesis of Dublin Core have also discussed.
- What is Dublin Core Metadata Standard and the Mission/Goal have been clearly described.
- Two Types, namely Simple Dublin Core and Qualified Dublin Core have been stated.
- 15 elements of Simple Dublin Core has been explained elaborately giving appropriate explanations as and when necessary. Further 3 more elements besides the above 15 have also mentioned.

4.3 METADATA

4.3.1 MEANING OF METADATA

- ‘Meta’ means something of a higher or more fundamental kind.
- Data which describes a resource (s) or data which is associated with an object that describes that object.
- Refers to networked electronic resources, all digital object that exists on the Internet and descriptive information about www.

- Also makes it easier to retrieve, use or manage an information resource.
- Commonly stored in database systems and linked to the objects described.
- HTML web page and the indexing data held by search engines is also metadata linked to the Information resource through Hypertext links to the URL.
- On the whole, the metadata can answer what, why, how, when, who and how about every facet of the data that are being documented.
- Metadata could be both Machine and Human generated description of a resource e.g. Bibliographic Database and OPAC.

4.3.2 PURPOSE OF METADATA

- To facilitate discovery of relevant information in the context of www and internet and helps to organize electronic resources.
- To be a source of bibliographic data.
- To facilitate users for data searching and browsing and instruct how to interpret the data and, how to transfer the data and data documentation.
- Facilitate inter operability and supports digital identification and support archiving and preservation
- To characterize the original work sufficiently for the user to understand its contents ,purpose and conditions of use.
- Serves as a basic tool for cataloguing of web resources and to organize vast amount of information available on the INTERNET
- For easy accessibility and effective use of the required information.

4.3.3 FUNCTIONS OF METADATA

Metadata describes the content, physical description, location, type and form of the information necessary for management which includes migration, history dates, security, authentication, file formats. The different functions of Metadata can be enumerated as under:

i. Resource Discovery :

- Metadata performs the same functions as an efficient, comprehensive and updated catalogue in resource discovery.
- Allowing resources to be found by relevant criteria.
- Identifying resources.

- Bringing similar resources together and separating dissimilar resources providing location information.

ii. **Organizing Electronic Resources**

Unprecedented growth of Web-based resources, aggregate sites/ portals is immensely useful in:

- Building the static Web Pages dynamically from meta data stored in data bases.
- Organizing Web Information through channels as channels are pre-selected Web sites that automatically “push” streams of information to a users browser used for continuously updated information.

iii. **Interoperability:**

- Interoperability is the ability of multiple systems with different hardware and software platforms, data structure, interface and Networks effectively to exchange with minimum loss of contents and functionality and to use the information that has been exchanged in a purposeful manner.

iv. **Digital Identification:**

Most of the metadata schemes include:

- ISBN to uniquely identify the work or object to which the metadata refers.
- The location of a digital object may also be assigned using a file name like Uniform Resource Locator (URL) or Digital Object Identifier (DOI). Persistent URL are preferable because file locations change frequently making URL and therefore metadata records invalid.

v. **Archiving and Preservation:**

As and when the hardware and software technologies change, the Digital resources will not survive in usable form in future because the digital information is fragile and every possibility of being corrupted or altered. Hence Metadata is:

- To ensure that resources will survive and continue to be accessible in the future.

- To track the lineage of a digital object (where it came and how it has changed over time) to detail its physical characteristics and to document its behavior in order to emulate it on future technologies, it is highly essential that Archiving preservation require special elements.

4.3.4 TYPES OF METADATA:

Meta data can be grouped into three types: Descriptive, Structural and Administrative.

i. Descriptive metadata

It refers to the intellectual content and associations of a document or resource in such a way that facilitates search identification and collection of information. It includes elements like title, author, date of creation or publication, abstract, keywords, information regarding the analog source from which a digital object is derived.

ii. Structural Metadata:

It describes the structure of file, dataset or other information package that is being described. Also indicates how related files are bound together and how the object can be displayed and disseminated on variety of systems. Narrates the attributes of an object like size, electronic format and digital capture process.

iii. Administrative Metadata:

It includes metadata for object management, right management, maintenance and preservation.

4.3.5 BENEFITS OF METADATA:

- Resource Discovery, Location and Resource Documentation
- Resource Selection, Evaluation and Assessment
- Safeguarding Intellectual Property Right
- Facilitates to filter obsolete data and provide only current information
- Permits to sort the information according to various forms besides according to subject content.

4.4 DUBLIN CORE

4.4.1 INTRODUCTION

Finding relevant information on the World Wide Web has become increasingly problematic due to unprecedented emergence of networked resources. Current web indexing to satisfy the demand for resource discovery tools is to be a poor substitute for richer varieties of resource description. Dublin Core is an initiative to create a digital “library card catalogue” for the web. The Dublin Core Metadata Element (DCME) Set is a standard for cross domain information resource description. It was developed to be simple and concise for describing Web-based documents. It is extensively used to describe digital materials such as composite media like Web pages, Image, Sound, Text and Video.

4.4.2 BACKGROUND:

The Dublin Core Metadata Initiative began in March 1995. The DCME Set arose from discussions at a Work shop in 1995 sponsored by OCLC and the National Centre for Super Computing Application (NCSA) which brought together librarians, digital library researchers, content experts and text mark-up experts to promote better discovery standards for electronic resources. As the work shop was convened in Dublin, Ohio, the element set was named as Dublin Core. The “Core” refers to the fact that the metadata elements set is a basic but expandable. The Dublin Core Metadata Initiative (DCMI) is looking after the continuing development of Dublin Core and related specifications and also took initiatives to promote wide spread acceptance of metadata standards and practices. Dublin Core is made up of 15 metadata elements that has emerged from this effort in interdisciplinary and international consensus building that offer expanded cataloguing information and improved document indexing for search engines programs.

4.4.3 WHAT IS DUBLIN CORE METADATA STANDARD?

- ❖ It is by far the most important metadata initiatives for web resource description.
- ❖ It is an exchange format for sharing records.
- ❖ Dublin core standard was intended to be descriptive rather than evaluating.
- ❖ It is an effective element set for describing a wide range of networked services.

- ❖ It is an excellent tool for resource discovery for networked resources.

4.44 MISSION OF DUBLIN CORE

- Confronted with exponential growth of electronic resources and problem to catalogue all these resources, the goal was:
- To define a few elements and some simple rules that could be applied by non cataloguers.
- Simplicity of creation and maintenance.
- Interoperability among collections and indexing systems.
- Extensibility
- Commonly understood semantics

4.45 FORMS OF DUBLIN CORE

Two forms of Dublin Core exists: 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 Elements Set(DCMES).

Qualified Dublin Core using Audience, Provenance and Rights Holder increases the specificity of metadata by adding information about encoding schemes, enumerated lists of values, or other processing clues.

4.46 THE DUBLIN CORE 15 ELEMENT SET

1. **TITLE.** The name given to the resource by the Creator or Publisher. A Title will be a name by which the resource is formally known. If the Title is in HTML then the source document be viewed and the title heard (if any) is also included as a Title.
2. **AUTHOR or CREATOR.** The person(s) or organization(s) primarily responsible for the intellectual content of the resource. For example, authors in the case of written documents, artists, photographers or illustrators in case of visual resources.
3. **SUBJECT or KEYWORDS.** The topic of the content of the resource. Typically, a Subject will be expressed as keywords, phrases, or classification codes that describe the topic of the resource.
4. **DESCRIPTION.** An account of textual description of the content of the resource, including abstracts in the case of document-like objects or content description in the case of visual resources and reference to a graphical representation of content.
5. **PUBLISHER.** The entity responsible for making the resource available in its present form, such as a publisher, a university department, or a corporate entity.

6. **OTHER CONTRIBUTORS.** Person(s) or organization(s) in addition to those specified in the Creator element who have made significant intellectual contributions to the resource, but whose contribution is secondary to the individuals or entities specified in the Creator element.
7. **DATE.** The date when the resource was made available in its present form. A date associated with an event in the life cycle of the resource.
8. **RESOURCE TYPE.** The category of the resource, such as home page, novel, poem, working paper, technical report, essay, dictionary. It is expected that Resource Type will be chosen from enumerated list of types.
9. **FORMAT. The physical or digital manifestation of resource.** The data representation of the resource, such as text/html, ASCII, Postscript file, executable application. Format will be assigned from enumerated lists such as Registered Internet Media Types. Format may include the media-type or dimensions of the resource.
10. **RESOURCE IDENTIFIER.** String or number used to uniquely identify the resource. Examples from networked resources include URLs , URNs. , Digital Identifier Object and ISBN.
11. **SOURCE.** The work, either print or electronic, from which this resource is delivered, if applicable. A Reference to a resource from which the present resource is derived.
12. **LANGUAGE.** Language(s) of the intellectual content of the resource.
13. **RELATION.** Relationship to other resources i.e reference to a related resource.
14. **COVERAGE.** The extent or scope of the content of the resource. The spatial locations(a place name or geographic co-ordinates) and temporal durations (a period label, date or date range) characteristics of the resource. .
15. **RIGHTS MANAGEMENT.** The content of this element is intended to be a link (a URL or other suitable URI as appropriate) to a copyright notice, a rights-management statement, or perhaps a server that would provide such information in a dynamic way. Rights information often encompasses Intellectual Property Rights, Copyright and various property rights.

4.47 QUALIFIED DUBLIN CORE.

- Audience, Provenance and Rights Holder are its elements but not part of Simple Dublin Core which has 15 elements.
- **AUDIENCE.** A class of entity for whom the resource is intended may be determined by the creator or publisher.
- **PROVENANCE.** A statement may include a description of any changes in ownership and custody of the resource since its creation that are significant for its authenticity, integrity and interpretation.
- **RIGHTS HOLDER.** A person or organization owning or managing rights over the resource. To indicate the entity it is better to use the URI or name of the Rights Holder.