Chapter XII
Metadata Interoperability

K. S. Chudamani
JRDTML, IISc, Bangalore-12, India

H. C. Nagarathna
JRDTML, IISc, Bangalore-12, India

ABSTRACT

Metadata is data about data. Metadata originated in the context of digital information in databases. This chapter looks at the various standards available for digital document description in the context of bibliographic databases. It also describes the variety of metadata associated with such systems. Some of the metadata standards examined are MARC21, Dublin Core (DC), and Libsys. The second part examines metadata interoperability and mapping among these standards.

INTRODUCTION

The first use of "**metadata**" originated in contexts related to digital information (chiefly with regard to databases). Since then, the general understanding of the term has broadened to include any kind of standardized descriptive information about resources, including nondigital ones. For example, library catalogues, abstracting and indexing services, archival finding aids, and museum documentation might all be seen as stored and retrieved based on metadata. The advantages of this are two fold. Firstly, it allows librarians, archivists, and museum documentation specialists to cooperate usefully across professional boundaries. Secondly, it enables the cultural heritage professions to communicate more effectively with those domains that also have an interest in metadata (e.g., software developers, publishers, the recording industry, television companies, the producers of digital educational content, and those concerned with geographical and satellite-based information). Therefore, metadata is critical to physical and intellectual accessibility and utility of digital document. In this sense, to quote Gilliland Swetland (2000),[5] "Metadata provides us with the Rosetta stone that will make possible to decode information objects and their transformation into knowledge in the cultural heritage information systems of the twenty first century." According to

Day, metadata is defined literally as "data about data." The term is normally understood to mean structured data about resources that can be used to help support a wide range of operations. These might include, for example, resource description and discovery, and the management of information resources and their long-term preservation.

Interoperable systems allow the exchange of information and sharing of resources. They focus on the storage of data in a **standard format**. The role of metadata in facilitating interoperability can be seen in number of contexts, such as:

1. e-governance;
2. electronic record management;
3. educational technology; and
4. library management systems.

Information about a resource needs to be embedded in the database. The catalogue record used for library management is the basis for identifying individual items and their management. Machine readable cataloguing **(MARC21)** covers all kinds of library materials and is used in automated library management systems in the Western libraries. Although most library management systems can import and export data in MARC21 (http://www.loc.gov/marc/bibliographic/ebcdmain.html) format, they usually have their own internal metadata standards. The availability of MARC records stimulated the development of searchable electronic catalogues. The user benefited from wider access to searchable catalogues, and later to a union catalogue, which allowed them to search several library catalogues at once.

Structured information is used to describe, explain, and locate resources. According to Harrasowitz's home page on e-journal resource guide,[5] "metadata functionality goes beyond the cataloguing functions of description and access to include content rating for filtering out sensitive or objectionable materials, the linking of physically separate information objects, and description of intellectual property rights of electronic publications."

Metadata used with one application can be used for several different purposes. Metadata in a library catalogue can be used to provide a variety of search options for the retrieval of items. Now-a-days, its main focus is on information retrieval and improving retrieval on the Internet. The Dublin Core (http://www.dublincore.org/) is currently based on a set of 15 data elements and can be used as a markup language in hypertext markup language (HTML), extensible markup language (XML), or RDF.

Metadata is required to allow users to search on a number of fields, such as author, title, and description. These facilities are useful not only for end users but also to the editorial team in order to collate statistics needed to generate reports. Storing metadata in a certain format and then converting it into HTML <META> tags using programs or scripts is meant to make the metadata conversion into other formats (such as XHTML) by altering the script.

Metadata is key in making resources accessible in the future. There are three main types of metadata:[4]

1. **Descriptive metadata:** These metadata describe a resource for purposes such as discovery and identification. They can include elements such as title, abstract, author, and keywords
2. **Structural metadata:** This type of metadata indicate how compound objects are put together, for example, how pages are ordered to form chapters
3. **Administrative metadata:** The administrative metadata gather information to help manage a resource, such as when and how it was created, file type, and other technical information. It also collects copyright information, such as who can accesses it. There are several subsets of cataloguers who make decisions about whether a catalogue record should be created for a whole set of volumes or for each particular volume in the