Article can be accessed online at http://www.publishingindia.com

Metadata Schemas and Cultural Resources: An Overview

Humma*, Tariq Ahmad Shah**

Abstract

The study is based on the content analysis of the existing literature on metadata and its importance with regard to its role in description and discovery of resources especially heritage objects. Cultural heritage objects which are remnants of civilizations and only link to get connected with our past need to be preserved carefully, promoted wisely and in cyberspace one has to ensure their increased accessibility. Metadata schema which provide essential links to and add visibility to cultural objects need to be framed thoroughly keeping intrinsic and extrinsic features of cultural objectives in consideration. The issues of interoperability between different metadata schemas need to be achieved with the implementation of effective metadata mapping and crosswalks. The current study provides a superficial view of such studies which have deliberated thoroughly on metadata and its importance to cultural objects when canvassed on digital frame.

Keywords: Metadata, Metadata Schemas, Cultural Objects, Controlled Vocabulary, Cross walksm Metadata Mapping.

1. Introduction

The term "*metadata*" was coined by Jack E. Myers in 1969, and it was first printed in a product brochure in 1973. Meyer used the term to represent current and future products associated with his *MetaModel* and later the term have been adopted by the computer science, statistical and library and information science communities to mean "data about data" (Greenberg, 2005). Metadata as defined by Miller is "the extra baggage associated with any resource that enables a real or potential user to find that resource; to decide whether or not it is of value to them; to discover where, when, and by whom it was created, as well as for what purpose; to know what tools will be needed to manipulate the resource; to determine whether or not they will actually be allowed to access to the resource itself and how much this will cost them" (as cited in Cwiok, 2005).

2. Taxonomy of Metadata Standards

The backbone of a metadata standard consists of a set of elements. The terms "schema", "scheme", and "element set" have been used interchangeably to refer metadata standards (Chan & Zeng, 2006). Zeng and Qin (2008) proposed a fourfold typology for designing a metadata schema which includes:

1. **Data** structure standards: are simply known as metadata standards. The most well known and widely used is Dublin core metadata element set (DCMES).
2. **Data content Standards**: These are rules / guidelines for entering value into elements. Example includes Anglo-American Cataloguing Rules II (AACR2) and Cataloguing Cultural Objects (CCO). Where ACCR2 is designed primarily for Monographs & Serials and CCO is designed particularly for cultural heritage communities.
3. **Standards for** data values: these include controlled term lists, classification schemas, thesauri and list of subject headings. Examples include Getty museums Art and architecture thesaurus (AAT), library of congress subject heading lists etc.

* Research Scholar, Department of Library & Information Science, University of Kashmir, Jammu & Kashmir, India. E-mail: humma.world@gmail.com

** Research Scholar, Department of Library & Information Science, University of Kashmir, Jammu & Kashmir, India. E-mail: traiqahmadshah@gmail.com

4. **Data** exchange /data format standards: these are required to encode metadata so that it can be easily processed by computer and exchanged among diverse systems. Extensible Markup language (XML) is the most widely used data exchange standard.

The above mentioned components are essential for the durability of any metadata schema.

The literature dealing with metadata schemas and their applications is growing. An overview of such studies particularly those concerned with heritage resources are deliberated under following sections:

3. Importance of Metadata

The significance of digital information depends upon how ably it can be located, searched for relevance and retrieval. With the advent of technology, organisation and retrieval have become more difficult. Finding items of interest on the web is increasingly becoming challenging because of the enormous growth of information resources, and therefore metadata has also become important to ensure that this growth of information objects can be organised and retrieved (Plessis, 2008). Metadata acts a key to ensure that resources will survive and continue to be accessible in future (Day, 2001). Sherbini and Klim (2004) observe that great proliferation of electronic content that is accessible through the World Wide Web makes it necessary to create and apply metadata standards that will make navigation more effectual. In the digital environment, the surrogate role of metadata is imperative because many resources are not easily browsable and others do not carry clear data about themselves (Coyle, 2005). Metadata is very useful in organizing chaotic world of digital resources and fundamental to effective retrieval and use of digital materials, as well as the administration, dissemination, and preservation of those materials in the online environment. Furthermore, metadata allows machine readability and understand ability, which is basically the vision for semantic web (Dhingra & Bhatia 2012; Good, 2010; Sherbini, 2001; Wisser, 2003). Pal (2010) also supports the view that the prime function of metadata creation is an effective strategy to enhance resource discovery from a large digital collection and makes the resources available to end users. In addition, it assures rights management, links to e-resources, enables interoperability using standard schemas and protocols (e.g., cross search by Z39.50 protocol or metadata harvesting using OAI protocol), Digital object identifier (DOI) and digital preservation. Acknowledging the importance of metadata for managing internet resources, Nikam (2002) concludes that effective management of networked information will definitely depend on the proper management of metadata. CDP Metadata Working Group (2008) believes that the creation of metadata for digital resources is an important part of a digitization project, and must be incorporated into a project's workflow. Metadata should be created and associated with a digital resource to support the discovery, use, management, reusability, and sustainability of the resource. Kenro et al (2007) ascertain that metadata can improve the search effectiveness and usability of the search system by providing multiple access points and preserving the semantics and context of the objects.

4. Metadata and Cultural Heritage Objects

The significance of cultural and heritage resources is immeasurable. Such resources often reflect the civilisation and culture of a nation and its people. Libraries, archives, historical societies and museums commonly referred to as cultural heritage institutions play an invaluable role in storing and preserving cultural heritage and making it available for use to the community. Cultural resources have long been accessible within the institutional boundaries. However, technological advancements have offered better opportunities not only to safeguard the cultural wealth, which was earlier limited to in-house use only but also provided a platform to showcase these resources to anyone who get connect to the virtual world. A number of cultural institutions are on the move to make their collections available online. Today, we are at a distance of a mouse click to know about the traditions and cultural behaviour of geographically dispersed people.

In the domain of digital cultural collections the concept of metadata is vital. While discussing the metadata schema for Japan's Cultural Heritage Online Project, Kando and Adachi (2004) annotate that key role of the metadata for cultural heritage is to provide access across heterogeneous objects. Heterogeneity is one of the main characteristics of cultural heritage collections. Cultural artefacts are usually complex, having being created by different cultures over different periods of time. Such collections may be composed of text written on different

materials, paintings, photographs, 3D objects, sound recordings, maps or even digital objects. Furthermore, the objects are strongly related with the social and historical events that take place over time. Consequently, it is quite justifiable to expose the composite structure, diverse semantics and multiple kinds of relationships between the objects of these collections. Hence, making accessible cultural heritage resources require metadata schemas rich in semantics and structure able to cover the material heterogeneity and variety of memory institutions (Kakali et al, 2007). Digital information systems for museum and archival collections make it easier to disseminate digital versions of unique objects to users around the globe who, for reasons of geography, economics, or other barriers, might otherwise not have an opportunity to view them. With new communities of users, however, come new challenges concerning how to make the materials most intellectually accessible. Cultural heritage information professionals such as museum registrars, library cataloguers, and archival processors often apply the term metadata to the value-added information that they create to arrange, describe, track, and otherwise enhance access to information objects and the physical collections related to those objects. Such metadata is frequently governed by community-developed and community-fostered standards and best practices in order to ensure quality, consistency, and interoperability (Gilliland, 2008).

5. Metadata Schemas and Cultural Heritage Objects

The rapid increase of Internet resources and digital collections has been accompanied by a proliferation of metadata schemas, each of which has been designed based on the requirements of particular user community, intended users, types of materials, subject domains, project needs, etc (Chan & Zeng, 2006). Metadata standards are designed to ensure data that is viable and can travel across disparate networks via a variety of hardware and software platforms without losing context or meaning in the process (Eklund, 2011). Several metadata schemas are being developed in a variety of user environments and disciplines. Many of the schemas have been adopted by many cultural institutions, among them are *Dublin Core, Metadata Object Description Schema (MODS), Encoded Archival Description (EAD), Visual Resources Association Core categories (VRA core), MPEG Multimedia Metadata* etc. (NISO, 2004). Besides these schemas there also exist metadata schemas for museums such as *Categories for the Description of Works of Art (CDWA)* (Lourdi & Nikolaidou, 2009). As far as the standardized metadata for documenting 3D content in the cultural heritage realm is considered, CARARE metadata schema has paid good attention. It provides rich set of classes with higher level of granularity needed for the description of 3D objects (Ronzino, Hermon & Niccolucci, 2012).

6. Data Value Standards / Controlled Vocabularies/ Thesauri

Cultural heritage institutions/repositories are bestowed with the responsibility of improving access to their cultural resources and for enhanced access concept of controlled vocabularies are crucial. Controlled vocabularies offer an important function in user discovery of digital resources and allows user to retrieve sets of meaningful related resources (Miller, 2011). Patra (2008) reveals that good metadata standards uses standard controlled vocabularies to reflect the what, where, when and who of the content. Moreover, if multiple organizations describe their collections consistently by using terms from a controlled list, the common approach will reap great benefits during searches. A wide range of controlled vocabularies are created and maintained by research institutions, national and international cultural organizations, professional societies and associations. They can be used individually or together, depending on the type of material being described (Harpring, 2010). The Getty vocabularies like *Art and Architecture Thesaurus (AAT), The Getty Thesaurus of Geographic Names (TGN), the Union List of Artist Names (ULAN), and the Cultural Objects Name Authority (CONA)* are intended to offer improved access to information about art, architecture and material culture (Getty Vocabularies, n.d). Other examples of controlled vocabularies used in cultural heritage metadata sector include Library of Congress Subject Headings, Thesaurus for graphic materials and the moving image genre. The *Thesaurus for Graphic Materials (TGM)* includes more than 7,000 subject terms to index topics shown or reflected in pictures, and 650 genre/format terms to index types of photographs, prints, design drawings, ephemera and other categories (About, n.d).

7. Crosswalks/Metadata Mapping

One of the most common challenges for cultural heritage institutions is to manage heterogeneous objects as a homogenous set of objects for reasons such as information retrieval and data exchange among diverse systems. Therefore, metadata interoperability mechanisms must be well-designed and applied (Bountouri & Gergatsoulis, 2009). A widely used method for metadata interoperability is *crosswalks* between various metadata schemas (Lourdi & Papatheodorou, 2008). As the number, size, and complexity of metadata standards continues to grow, supplying the metadata for each standard becomes more and more time-consuming and tedious. In order to minimize the amount of time needed to create and maintain the metadata and to maximize its usefulness to the widest community of users, there is a need for the metadata created and maintained in one standard to be accessible via related content metadata standards (Pierre & LaPlant, Jr., 1998). Metadata mapping is the process of identifying equivalent or nearly equivalent metadata elements or groups of metadata elements within different metadata schemas, carried out in order to facilitate semantic interoperability. The users can search for digital information across heterogeneous distributed databases as if they were all part of the same virtual repository (Baca, 2003). As mentioned by Woodley (2008), metadata mapping refers to an intellectual process of comparing and examining two or more metadata standards and then determining which elements in one schema maps to elements in the other. While as crosswalks are the visual and textual product (usually in form of tables) of the mapping process. Over the years a good number of crosswalks have evolved though the exact list of the tables is not documented officially anywhere. Some of the metadata standard cross walks are: MODS to Dublin core (unqualified); VRA Core to Dublin core (unqualified); MARC to MODS; CDWA to CDWA lite, etc. (http://www.getty.edu/research/publications/electronic_publications/intrometadata/crosswalks.pdf)

8. Conclusion

Metadata described as a key to resource discovery and access plays an important part in the optimum utilization of resources. Unless digital resources are not properly describes with appropriate standards there is very likelihood that they will get lost in the ocean of World Wide Web and the essence for which they are hosted online will be forfeited. Cultural resources which report a link to civilizations and are an important facet of our identity need to be dealt with great care. In the current digital environment the job of cultural heritage institutions do not end with the digitization of their resources but, the organisations need to guarantee an effective link between the resource and its potential users. And the first step to ensure proper linkage comes with the adoption and implementation of most suitable metadata schema. Since cultural objects represent a much diverse and heterogeneous group of resources, it is imperative to select suitable schema which will illustrate the resource at more granular level. Further, there is need to have control over the terminology being used to describe various facets of heritage objects. And to achieve the same implementation of appropriate control vocabulary/thesauri is necessity. As commented by Baca (2003), the dream of integrated access to diverse information resources can become a reality if those responsible for making cultural heritage information available online judiciously select and implement the appropriate metadata schemas, controlled subject vocabularies and thesauri, metadata crosswalks, and information technologies available to us today.

References

[1] About the Thesaurus for Graphic Materials. Washington DC: Library of Congress. Retrieved from http://www.loc.gov/pictures/collection/tgm/ (accessed on October 22, 2010)

[2] Alemu, G., Stevens, B. & Ross, P. (2012). Towards a conceptual framework for user-driven semantic metadata interoperability in digital libraries: A social constructivist approach. *New Library World*, 113(1-2), 38-54. DOI: 10.1108/03074801211199031

[3] Baca, M. (2003). Practical issues in applying metadata schemas and controlled vocabularies to cultural heritage information. *Cataloging & Classification Quarterly*, 36(3-4), 47-55. DOI: 10.1300/J104v36n03_05

[4] Bountouri, L. & Gergatsoulis, M. (2009). Interoperability between archival and bibliographic metadata: An EAD to MODS crosswalk. *Journal of Library Metadata*, 9(1), 98-133. DOI: 10.1080/19386380903095107

[5] CDP Metadata Working Group. (2008). *Dublin Core Metadata: Best Practices*. Retrieved from http://www.mndigital.org/digitizing/standards/metadata.pdf

[6] Chan, L. M. & Zeng, M. L. (2006). Metadata interoperability and standardization: A study of Methodology Part I: Achieving Interoperability at the Schema Level. *D-Lib Magazine,* 12(6). Retrieved from http://dlib.org/dlib/june06/chan/06chan.html

[7] Coyle, K. (2005). Understanding metadata and its purpose. *Journal of Academic Librarianship,* 31(2), 160-163. DOI: 10.1016/j.acalib.2004.12.010

[8] Cwiok, J. (2005). The defining element- A discussion of the creator element within metadata schemas. *Cataloguing & Classification Quarterly,* 40(3-4), 103-133. DOI: 10.1300/J104v40n03_06

[9] Day, M. (2001). *Metadata in a nutshell*. University of Bath, UK: UKOLN: the UK Office for Library and Information Networking. Retrieved from http://www.ukoln.ac.uk/metadata/publications/nutshell/

[10] Dhingra, V., & Bhatia, K.K. (2012). Metadata towards machine enabled intelligence, *International Journal of Web and Semantic Technology,* 3 (3), 121-130. DOI: 10.5121/ijwest.2012.3308

[11] Eklund, J. L. (2011). Cultural objects digitization planning: Metadata overview. *VRA Bulletin,* 38(1), 1-19. Retrieved from http://online.vraweb.org/vrab/vol38/iss1/4

[12] Getty Vocabularies. (n.d). The Getty Research Institute. Retrieved from http://www.getty.edu/research/tools/vocabularies/index.html

[13] Gilliland, A. J. (2008). Setting the Stage. In Murtha Baca (ed.), *Introduction to Metadata.* Los Angeles: Getty Research Institute.

[14] Good, J. (2010). *A Gentle Introduction to Metadata.* Retrieved from http://openaccess.uoc.edu/webapps/o2/bitstream/10609/7601/7/Good_AGentleIntroductionToMetadata.pdf

[15] Greenberg, J. (2005). Understanding metadata and metadata schemes greenberg. *Cataloging & Classification Quarterly*, 40(3-4), 17-36. DOI: 10.1300/j104v40n03_02

[16] Harpring, P. (2010). Vocabularies for Cultural Objects. *In* Murtha Baca (Series Ed.), *Introduction to Controlled Vocabularies: Terminology for Art, Architecture, and Other Cultural Works*. California: Getty Publications. Retrieved from http://www.getty.edu/research/publications/electronic_publications/intro_controlled_vocab/cultural_objects.html

[17] Kakali, C., Lourdi, I., Stasinopoulou, T., Bountouri, L., Papatheodorou,C., Doer, M. & Gergatsoulis, M. (2003). Integrating Dublin Core metadata for cultural heritage collections using ontologies. *In International conference on Dublin Core and Metadata Applications profiles: Theory and Practice*, August 27-31, pp. 128-138. Singapore. Retrieved from dcpapers.dublincore.org/index.php/pubs/article/view/871/867

[18] Kando, N. & Adachi, J. (2004). *Cultural Heritage Online: Information Access across Heterogeneous Cultural Heritage in Japan*. In International Symposium on Digital Libraries and Knowledge Communities in Networked Information Society (DLKC'04), March, 2(5), 136. University of Tsukuba, Japan. Retrieved from http://www.kc.tsukuba.ac.jp/dlkc/e-proceedings/papers/dlkc04pp136.pdf

[19] Kenro, A. (2007). Building a Space of Cultural Heritage Objects to explore it in the Classroom. In Proceedings of the International Conference of ICL2007, September 26 -28, pp.1-18. Villach, Austria. Retrieved from http://hal.inria.fr/docs/00/25/71/15/PDF/189_Final_Paper.pdf

[20] Lourdi, E. & Nikolaidou, M. (2009). Guidelines for Developing Digital Cultural Collections. In Theng, Y., Foo, S., Goh, H. L. D. & Na, J. C. (Eds.), *Handbook of Research on Digital libraries: Design, Development and Impact* (pp. 198-205). Hershey, P.A: IGI Global.

[21] Lourdi, I. & Papatheodorou, C. (2008). *Semantic Integration of Collection-Level Information: a crosswalk between CIDO/CRM & Dublin Core Collections Application Profile*. 2008 Annual Conference of CIDOC, Athens, September 15-18, 2008. Retrieved from http://www.ionio.gr/~papatheodor/papers/cidoc2008.pdf

[22] Miller, S. J. (2011). *Metadata for Digital Collections: A how-to-Do-it Manual*. London: Facet Publishing.

[23] Nikam, K. (2002). *Metadata for Managing Internet Resources*. Workshop on Information Resource Management (pp. 1-8). DRTC, Bangalore. Retrieved from http://drtc.isibang.ac.in:8080/bitstream/handle/1849/82/Paper-CH.PDF?sequence=2

[24] NISO (National Information Standards Organization). (2004). Understanding Metadata. Bethesda, MD: NISO Press. Retrieved from www.niso.org/publications/press/understandingMetadata.pdf

[25] Pal, J. K. (2010). Metadata initiatives and emerging technologies to improve resource discovery. *Annals of Library and Information Studies*, March, 57(1), 44-53. Retrieved from http://nopr.niscair.res.in/bitstream/123456789/8284/4/ALIS%2057(1)%2044- 53.pdf

[26] Patra, C. (2008). Digital repository in ceramics: A metadata study. *The Electronic Library*, 26(4), 561-581. DOI: 10.1108/02640470810893792

[27] Pierre, M. S. & LaPlant, Jr., W. P. (1998). *Issues in Crosswalking Content Metadata Standards, NISO*. Retrieved from http://www.niso.org/publications/white_papers/crosswalk/

[28] Plessis, M. D. (2008). Metadata for Digital Resources: Implementation, Systems Design and Interoperability. *The Electronic Library*, 26(6), 925-926.

[29] Ronzino, P., Hermon, S. & Niccolucci, F. I. (2012). A Metadata Schema for Cultural Heritage documentation. Retrieved from http://www.academia.edu/1227094/A_Metadata_Schema_for_Cultural_Heritage_Documentation

[30] Sherbini, M. E. & Klim, G. (2004). Metadata and cataloguing practices. *The Electronic Library*, 22(3), 238-248. DOI: 10.1108/02640470410541633

[31] Sherbini, M. E. (2001). Metadata and the future of cataloguing. *Library Review*, 50(1), 16-27. DOI: 10.1108/00242530110363217

[32] Wisser, K. (2005). Meeting metadata challenges in the consortial environment: Metadata coordination for North Carolina Exploring Cultural Heritage Online. *Library Hi Tech*, 23(2), 164-171. DOI: 10.1108/07378830510605133

[33] Woodley, M. S. (2008). Crosswalks, Metadata Harvesting, Federated Searching, Metasearching: Using Metadata to Connect Users and Information. *In* Murtha Baca (Ed.), *Introduction to Metadata* (pp. 38-62). Los Angeles: Getty Information Institute. Retrieved from http://nsdl.niscair.res.in/bitstream/123456789/954/1/Introduction+to+Metadata.pdf

[34] Zeng, M. L. & Qin, J. (2008). *Metadata (pp. 15-16)*. London: Facet Publishing