

Our Visual Cultural Heritage Storage: Introduction to MIX Metadata Standard for Visual Materials

Görsel Kültürel Mirasımızın Depolanması: Görsel Materyaller İçin MIX Üst Veri Standardına Giriş

Solmaz Zardary

University of Tehran, LIS Department, Iran. szardary@gmail.com

Fatima Fahimnia

University of Tehran, LIS Department, Iran. fahimnia@ut.ac.ir

Abstract: *Our visual cultural heritage is one of the important parts of contemporary digital storages and libraries. In this essay, using a special kind of metadata for storage of images in "digital repositories" is reported. The National Information Standards Organization (NISO) in its data dictionary mentioned as the goal of metadata standards to enable users to develop, exchange, and interpret digital resources. This paper briefly introduces a metadata standard developed by the Library of Congress for the storage of cultural heritage, MIX: Metadata for Digital still Images Standard in XML (eXtensible Markup Language) and reviews its features.*

Keywords: *Metadata, MIX, metadata standards, digital still images*

Öz: *Görsel kültürel mirasımız çağdaş dijital arşivlerin ve kütüphanelerin önemli parçalarından biridir. Bu denemede "dijital arşivler"de görüntülerin depolanması için özel bir üst veri türünün kullanımı rapor edilmektedir. Amerikan Ulusal Standartlar Örgütü NISO veri sözlüğünde üst veri standartlarının amacının kullanıcıların dijital kaynakları geliştirmelerine, değiştirmelerine ve yorumlamalarına olanak sağlamak olduğundan söz etmektedir. Bu bildiri Kongre Kütüphanesi tarafından kültürel mirasın depolanması için geliştirilen MIX (Dijital Hareketsiz Görüntüler İçin XML –Genişletilebilir İşaretleme Dili) üst veri standardını kısaca tanıtmakta ve özelliklerini incelemektedir.*

Anahtar sözcükler: *Üst veri, MIX, üst veri standartları, dijital hareketsiz görüntüler*

Introduction

In today's information society and the World Wide Web, the new generation of users expect libraries to be a treasury of knowledge to serve in their progress. They need libraries' informational support in new ways. Users access information from their offices, homes or other places without physical connection to libraries. They find their information via the internet. But the accuracy and reliability of this kind of information is questionable. It is the task of the library to collect, classify and disseminate information as a center of the information society in this user-oriented age. In order to classify this kind of information, many scholars, experts and organizations are making great efforts. Their aim is to have uniform and common principals in this sharable world so that they make information accessible for all over the world without limitations. The goal of standardization in virtual libraries is to insure interoperability between systems in long term management and accessibility (NISO, 2006). NISO (2006) defined management as tasks and operations needed to support image quality assessment, image data processing, and long term maintenance throughout the image life cycle.

Experts have also been trying to have this uniformity via standardized metadata in digital world. Commonly, metadata is defined as data about data. The most prevalent definition for metadata says that it is the structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use or manage an information resource. There are many standardized metadata that are used to make the digital library accessible, full of basic information, structural and sharable.

One kind of information derived from our ancestors is illustrations and images. In that special kind of communication, they transmitted their information. Even now sometimes we use this kind of communication. Some people believe one image is more communicative and clearer than thousands of words. Also we have a big amount of visual heritage.

Most libraries, archives, art galleries, publishers, rights holders, cultural heritage institutions, and storage centers have been saving this kind of information in a virtual environment. To standardize the storage process, they use

special metadata. In this essay we are trying to introduce one of the metadata standards that is used in these organizations to store the virtual information. It is important to mention that this standard is not complete or the only standard in this arena, because some other standards can also be used.

MIX

This technical descriptive metadata is for digital still images in XML environment and is presented by the Library of Congress Network Development and MARC Standards Office. It is supported by the National Information Standards Organization (NISO) Technical Metadata for Digital Still Images Standards Committee and other scholars (Waibel & Dale, 2004). Guenther (2007) describes technical metadata details like image capture information, bit depth, color space, photometric interpretation, and specific format characteristics. "The schema provides a format for interchange and/or storage of the data specified in the Data Dictionary - Technical Metadata for Digital Still Images (ANSI/NISO Z39.87-2006). This schema is currently referred to as 'NISO Metadata for Images in XML (NISO MIX)'. MIX is expressed using the XML schema language of the world wide web consortium". The MIX schema after emerging and trials has edited several times. A draft version 0.2 has been available since July 2004. 2.0 is the current version (Needleman, 2005). Access to its previous versions, documents, examples, tools, utilities and comments about this schema are available via the Library of Congress web site (Library of Congress, 2004). Galloway (2004) imparts: "The Library of Congress, motivated by its Digital Audio-Visual Preservation Prototyping Project created a metadata scheme based on the Z39.87 standard, called MIX". It is noteworthy that this schema is for the data elements of still images and that for motion images or other kind of images there are other kinds of standards that are not considered as the purpose of this article.

We can insert large varieties of data information "from basic information on file types and sizes, to details of image capture (including capture hardware and image targets), to details of how an image has been processed after capture. Although a MIX file can be very lengthy and complex, almost all of its components (more than in its parent element set) are optional so that a basic record may be very simple. Although still in the process of revision (version 2.0 is in draft at the time of writing), MIX has already established itself as the key standard for this type of metadata" (Gartner, 2008).

MIX is not tied to any specific format. Thus it covers all image formats of today's technology and future unknown formats. It contains data elements common to all image file formats (e.g. pixel dimensions). "It does not include formats because it is critical to long term preservation and plays a primary role in any preservation strategy" (NISO, 2006).

The MIX metadata set works with other existing standards and is usable and compatible in a variety of application domains and user situations. It can be integrated with utilization of XML schemas. For example MIX can be integrated with other metadata standards in XML such as PREMIS (Preservation Metadata Maintenance Activity) or METS (Metadata Encoding and Transmission Schema) as extension schema for technical metadata.

MIX also appeared in the data dictionary- technical metadata for digital still images published by NISO, 2006. All metadata and their usages and definitions that are required for digital still image are illustrated. But MIX is not from the main part of this dictionary. It is one part of an appendix. Some mentioned four parts for MIX (e.g. Morris, 2005; Caplan, 2003; Galloway, 2004):

Basic image parameters: They document the file structure with sufficient parameters to insure a reasonable image can be rendered for viewing, such as format information like MIME type, compression, and file information, such as file size, checksum, and orientation. These parameters are crucial to displaying a viewable image

Image creation: This information is crucial to understand the technical environment in which a digital image was captured; describes how, when and by whom the image was created. This metadata set records detailed information about image capture, whether this was done with a digital camera or by scanning from an analog source.

Image performance assessment: This element contains the metadata pertaining to image quality or output accuracy, such as color management, digitization resolution. This set is designed to serve as metrics to assess the accuracy of output (today's use), and to assess the accuracy of preservation techniques, particularly migration (future use). These metrics can include length and width of image or color components per pixel.

Change history: It supports tracking of image processing that has been applied to the image. It documents any editing operations performed on the image during its life cycle, including the responsible party, the date and time and the software used in the course of refreshing or migration.

Features of MIX

Some features of this metadata standard are mentioned below:

MIX is supported by a responsible, reliable institution and organization (the Library of Congress and NISO). MIX is based on the NISO metadata set. Thus it has interoperability, exchangeability, and usability potential in most situations.

The kind of details about an image which this metadata preserves is very useful information for people who work with storage and retrieval.

MIX is a composition of several metadata sets for storage (McCallum, 2005); it is possible to use some rudimentary parts of it and not all and vice versa.

Also Wilson and his teammates (2009, p. 15) recounted some other features for this metadata as follow:

- *A comprehensive standard developed specifically for the management of digital still images*
- *Extensible and scalable – additional fields can be created and used*
- *Not tied to any specific file format – Includes data elements common to all image file formats*
- *Able to work well with existing standards and integrate with XML schemas*
- *Able to stand alone or be used as part of a broader metadata set like PREMIS*

Discussion and Conclusion

Nowadays images are one of the visual sources that contain information. Lots of effort has gone into storage, maintenance, accessibility and retrieval with ultimate aim being to serve users. New technology and its use in the image arena has made digital image repositories in a focal point of experts. As a standard and uniform means of image storage the MIX metadata set is notable. MIX is not the totality of these repositories but it can help to develop huge amounts of our cultural heritage as images in the digital world.

References

- Caplan, P. (2003). *Metadata fundamentals for all librarians*. Chicago: American Library Association.
- Galloway, P. (2004). *Metadata requirements for a digital repository to accompany the American Anthropological Society's AnthroSource portal project*. Retrieved April 20, 2010, from dspace at the university of Texas at Austin school of information website <https://pacer.ischool.utexas.edu/bitstream/2081/54/1/metadata%20paper.pdf>
- Gartner, R. (2008). *Metadata for digital libraries: State of the art and future directions*. Retrieved April 17, 2010, from JISC: supporting education and research website http://www.jisc.ac.uk/media/documents/techwatch/tsw_0801pdf.pdf
- Guenther, R. (2007). MIX: what it stands for: Metadata for Images in XML Schema (Primers on standards). *Computers in Libraries*. 27(4), 16.
- Library of Congress. (2004). *MIX: NISO Metadata for image in XML Schema Technical Metadata for Digital still images Standard*. Retrieved April 20, 2010, from: <http://www.loc.gov/standards/mix/>
- McCallum, S. (2005). Preservation metadata standards for digital resources: What we have and what we need. *World Library and Information Congress: 71st IFLA general Conference and Council, 2005*. Retrieved 20 March 2010, from IFLA: the official website of the international federation of library associations and institutions <http://archive.ifla.org/IV/ifla71/papers/060e-McCallum.pdf>
- Morris, R.A. (2005). *Image metadata standards and practices*. Retrieved April 13, 2010, from GBIF CIRCA website <http://circa.gbif.net/irc/download>
- Needleman, M. (2005). Standards update: some interesting XML standards. *Serials Review*, 31(1), 70–71.
- NISO. (2006). *Data dictionary: Technical metadata for digital still images*. NISO Press: United State. Retrieved March 13, 2010, from National Information Standards Organization website http://www.niso.org/kst/reports/standards/kfile_download?id%3Aastring%3Aiso-8859-1=Z39-872006.pdf&pt=RkGKiXzW643YeUaYUqZ1BFwDhIG424RjbcZBWg8uE4vWdpZsJDs4RjLz0t90_d5_ymGsj_IKVa86hjP37r_hKQ00ioOP35W6Q1wkBOLKanPbfamndQa6zkS6rLL3oIr
- Waibel, G. & Dale, R. (2004). Automatic exposure: capturing technical metadata for digital still images. *Proceedings of IS&T 2004 Archiving Conference*. Retrieved April 15, 2010, from http://www.oclc.org/research/activities/past/rlg/automaticexposure/ae_whitepaper_2003.pdf
- Wilson, K., Billington, L., Moir, S., & Carpenter, S. (2007). Development of a metadata application profile at the State Library of New South Wales, *Proceedings of the 13th Information Online Conference*, Australian Library and Information Association, Sydney. Retrieved March 7, 2010, from http://www.sl.nsw.gov.au/about/publications/staffpapersdocs/map_paper.pdf