Andy Uhrich
Digital Preservation
2009-12-02

The Role of Metadata in the Preservation of Digital Audio Visual Content

It would be hard to overestimate the role of metadata in preserving digital moving image and sound content. While consistent metadata is absolutely crucial for exerting intellectual and physical control over analogue AV objects, their material presence confers varying degrees of human readability outside the directing lens of metadata. However, we are almost totally reliant on metadata in some form to *read* digital multimedia content, whether reading means playing it back, retrieving it, sharing it, searching it, or maintaining it over time. Somewhat similarly, while every point along the chain of digital preservation is equally important – including the codec and wrapper chosen for the file, the architecture of the file repository, the choice of collections management software, open source versus proprietary, storage in external hard drive versus LTO, the interface to provide access, etc. – metadata is the conceptual device that links all of these elements together.

Without properly conceived, gathered, and regulated metadata, digital AV content will be lost in the binary wilderness; underutilized, eventually overtaken by technological change, and a waste of precious institutional resources. As such, the importance of metadata in digital preservation is not simply limited to the subset of metadata fields concerned with preservation activities, however important those are, but should be thought of as a valuable tool for preservation

in its own right. Metadata is not just the aggregation of information about content, but is a set of rules for organizing, creating, and disseminating the information; and this dissemination and transfer of information is the core function of the digital environment/marketplace/library. Additionally, for the metadata to guarantee the functionality, retrievability, exchangeability, and viability of digital AV content over time the metadata itself must be functional, retrievable, exchangeable, and viable. So reified, metadata can be shaped by basic principles and best practices that ensure the tool is operated correctly.

Before moving on to those principles and practices a couple of establishing points are necessary. The first concerns the larger purpose of metadata regardless of format or analogue versus digital. According to the NISO's *Understanding Metadata* metadata is "structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource". It goes on to describe the three common subgroups of metadata and the domain they cover: descriptive, structural, and administrative[1]. While the NISO guide, like many others, places preservation metadata under the administrative category, as Brian Lavoie and Richard Garter point out "this categorization is not correct" as "a preservation metadata schema will include descriptive, structural, and administrative metadata elements"[2]. Therefore, the preservation of digital AV content requires all forms of metadata and not just a narrow subset called *preservation metadata*. In the analog realm

[1] National Information Standards Organization, *Understanding Metadata*, Bethesda: NISO Press, (2004): 1, http://www.niso.org/publications/press/UnderstandingMetadata.pdf.
[2] Lavoie, Brian and Richard Gartner, *Technology Watch Report: Preservation Metadata*, Digital Preservation Coalition, (2005): 4, http://www.dpconline.org/docs/reports/dpctw05-01.pdf.

preservation metadata can be more compartmentalized. It might, for example, cover the date that a damaged film print was repaired or transferred to a new film stock. As AV objects in the digital world are more mutable and ephemeral, to preserve the content's reference, context, provenance and fixity[3], metadata must combine information from the domain of the descriptive, structural, administrative, rights, and preservation.

In pivoting from what metadata is and what types of information it records to the principles and practices that govern its ideal formation, the use and meaning of the word *structured* in the NISO definition of metadata becomes paramount. The structure of the metadata chosen (or adapted or created) and the manner in which it is implemented determines how successful the metadata will be in preserving the content. In regards to the foundational principles underlying metadata, NISO's *A Framework of Guidance for Building Good Digital Collections*[4] submits six necessary principles that the metadata in good digital collections must follow[5].

The first principle is that metadata should adopt the standards of the institution's community, reflect the technical needs of the objects in the collection, and be organized in a manner that makes sense to its users. Different fields have

---

[3] According to the OAIS, these are the elements that the Preservation Description Information must record to preserve the Content Information. See page 28 of *Reference Model for an Open Archival Information System* at http://public.ccsds.org/publications/archive/650x0b1.pdf.

[4] Their description of what constitutes a good digital collection lies at the center of this discussion. As they discuss, the concept of a good digital collection has evolved over time with changes in technology and user habits. While good once meant faithfully serving a local and known audience, now good means that the digital collection must be usable by unexpected audiences for new purposes. This implicitly requires a more robust and responsive form of metadata at higher levels of precision and granularity.

[5] The six principles for metadata are found in: NISO Framework Working Group, *A Framework of Guidance for Building Good Digital Collections: 3rd Edition*, National Information Standards Organization, (2005): 63-85, http://www.niso.org/publications/rp/framework3.pdf.

a history of using different metadata structures and rules that reflect local needs: the library world developed MARC21 and AACR2 to catalogue books and the system is optimized for individual instances of mass produced titles; the traditional paper archive world has organized around Encoded Archival Description and DACS, which was conceptualized for describing collections at the series and the folder levels; public broadcasting stations refined and expanded Dublin Core to create PBCore; CDWA Lite works best with art works and objects of material culture; and so on. The NISO's *Framework* lists 19 metadata structures in use by cultural institutions. While this array of schemes can seem overwhelming and daunting, almost like a digital Tower of Babel, it allows for organizations to combine elements from each to suit the needs of their collection.

Additionally, while many metadata structures, such as MARC and EAD, were created to translate the library and archive's traditional methods of cataloguing/arrangement into the standardized and exchangeable digital realm, as they have started to collect objects outside of their traditional ambit they have had to adopt elements from metadata structures outside of their field. For example, when MIT's library created its digital repository, DSpace, they chose Dublin Core over MARC as it was more flexible for the variety of new digital formats they were collecting[6]. In many ways this is something that moving image and sound archives should be adept at as they have always been an ill fit with pre-existing standards such as MARC and EAD and have had to shift or adapt

[6] Lubas, Rebecca L., Robert H.W. Wolfe and Maximilian Fleischman, "Creating Metadata Practices for MIT's OpenCourseWare Project," *Library Hi Tech* 22.2 (2004): 138.

them to fit. It's important to note that any changes or additions to pre-existing metadata structures be well documented and fully explained by the institution.

Choosing the field's pre-existing metadata structure, if it's a logical fit, is beneficial as a means of using a structure that is widely adopted, regulated to some degree, and well documented. These aspects contribute to NISO's second principle for metadata: that it be interoperable. This is in many ways the core principle that all others are working to support. While they focus on a current form of interoperability – between different schema, across different platforms and for different types of users – it's crucial to add to this principle that the metadata must also be interoperable in the future for the preservation of the metadata through migration. There are a number of ways to facilitate the exchange of metadata, such as mapping, or crosswalking, between structures or being open to harvesting for federated databases such as through the Open Archives Initiative. The success rate of mapping relies on the granularity of the metadata and the degree to which the fields are documented. A finer granularity allows for information to be mapped to its precise analogue in the other structure as well as facilitating more accurate searches by users. Better documentation of the fields in a structure enables for a side-by-side of the different schemas. Just as important is the manner in which the fields are populated. It's vital that standardized semantic and syntactic rules and values direct the way information is entered into the metadata structure. This ensures that a search for a creator, subject, geographic space, etc. returns the desired hits since the names are

conformed to a name/subject authority, adopt a common spelling, and are entered in the same manner.

Which is exactly NISO's third principle: that good metadata employs aforementioned authority in the form of thesauri and controlled vocabularies. There are a number of widely accepted thesauri for various fields such as the Library of Congress Name and Subject Authorities, Getty's Art & Architecture Thesaurus, and the Thesaurus for Graphic Materials. As with combining fields from various metadata structures it is expected that institutions with digital moving image and sound collections will appropriate controlled vocabulary from established thesauri as is relevant and create their own to cover local names and subjects. Once again, for the sake of interoperability it is crucial that this process be sufficiently documented by noting from which thesauri the name authority comes from and for local names noting their creation and syntactical style followed.

The fourth principle concerns the importance of rights information. The speed and ease at which digital items can be distributed and discovered makes it essential that institutions have a clear understanding of the rights of their digital collection. While this subject is mostly outside of the boundaries of this paper, besides limiting interoperability, restricted rights can prevent preservation of content.

NISOs fifth and sixth principles are closely related and have already been mentioned in this paper. The fifth is that metadata should sustain "the long-term

management, curation and preservation of objects in collections"[7]. As previously discussed this involves a combination of structural, administrative, rights, and descriptive metadata – any information that allows AV content to be faithfully reproduced in a new digital environment. The sixth principle is that the metadata is a valuable digital object and subject to all of the considerations as digital information worth preserving.

The practices that achieve these principles revolve around consistency and accuracy. Consistency speaks to how the fields are structured, in how data is entered into them, in the methods of documenting their creation and use, and in training employees in how to record the metadata. Accuracy speaks to the degree to which the consistency is achieved.

Some of the following practices have already been mentioned but it's worth gathering them together here. First, a more robust metadata structure requires a greater outlay of funds, a higher level of skill and knowledge to set up, and a commitment to setting-aside larger amounts of future resources to properly manage. Therefore, when choosing or creating a metadata structure an organization needs to realistically base those choices on their available resources and limitations. A smaller number of fields of metadata with less granularity when consistently followed is more interoperable and adds preservation much better than an inaccurately populated larger and more granular structure. The scale of the metadata structure should mirror the scale of the organization.

[7] NISO, *Framework*, 83.

Towards this point, the confusing array of acronymed metadata structures is actually beneficial. Organizations can pick one that fits their resources and needs. Further, they can choose a schema that reflects the needs of the items in their collection and the level to which they are preserving them digitally, i.e. just the content, analogous to reformatting newspapers to microfilm, or the complete experience of the object. As Priscilla Caplan notes in a discussion of preserving text-based items "preservation metadata must define the essential characteristics of a record, and will vary according to the situation. If the content is essential, then the markup of that content can be considered preservation metadata. If the font is also essential, then preservation metadata must record the font"[8]. Obviously, with AV items it is more difficult to separate the content from the experience, but the point remains that the form of the metadata structure adopted should reflect the preservation requirements of the objects it will document.

As there is no one single structure that can be recommended as representing best practice, and to reflect both the local needs of an institution and the NISO principle of interoperability how should an organization construct their metadata schema? First, research for this paper indicates that the metadata for most digital collections combines fields from different pre-existing structures. In 2003 the Moving Image Collections at Rutgers University combined descriptive metadata from Dublin Core with technical metadata from MPEG-7[9]. Wilson et al in describing the minimum requirements for digital AV metadata combine fields

---

[8] Priscilla Caplan, "Preservation Metadata", *DCC Digital Curation Manual*, ed. S.Ross, M.Day, (July 2006):12, http://www.dcc.ac.uk/resource/curation-manual/chapters/preservation-metadata.

[9] Agnew, Grace and Dan Kniesner, *MIC – ViDe Application Profile and MS-Access Database Implementation for MPEG-7 and Dublin Core*, Rutgers University, (2003), http://gondolin.rutgers.edu/MIC/text/how/mpeg7_userGuide_ver_1.pdf.

from Dublin Core, PREMIS, MPEG-7, VideoMD, TV-Anytime, and AudioMD[10]. Ideally, all fields chosen will come from widely adopted and well-documented structures as they will already be described and mapped. Locally created fields should only be created when absolutely necessary and then documented and mapped.

What seems to be best practice now for the digital AV metadata world when creating an integrated metadata structure is to choose a schema as the core schema for the descriptive metadata and then pick and choose needed fields for technical, rights, and preservation metadata. The core schema could be Dublin Core, MODS, or PB Core. They all have the advantage of being easy to comprehend and map. Dublin Core in its stripping metadata down to the essential fields can be too simplistic and not robust enough, but with qualifiers can be more responsive. MODS is closely related to MARC so would be of use to a digital collection related to a library. Dublin Core was created to meet the needs of public broadcasting, but users outside of that precise world have begun to adopt it. Fields for preservation and some rights metadata can be gathered from PREMIS, though as Wilson et al note, it "is of little use in relation to the significant characteristics of audio and visual resources[11]. For these it make sense to turn to the more technically specific schemas such as SMPTE or MPEG-7 that might be difficult to use as the core schema for many users, but which describe the technical specifications of AV formats precisely and

---

[10] Wilson, Andrew et al, *Digital Moving Images and Sound Archiving Study*, Arts and Humanities Data Service, (2006): 83-89, http://www.jisc.ac.uk/media/documents/programmes/preservation/moving%20pictures%20and%20sound%20archiving%20final%20version.doc.

[11] Wilson, et al, 79.

comprehensively. As time-based items, digital video and audio is significantly more complex than still images and text and as such the importance of technical metadata should not be overlooked. Wilson et al's minimum requirements for AV metadata offers a useful place to start when deciding on what sort of technical metadata to capture[12].

A related metadata issue is whether or not to embed the metadata with the item or not, and if so which wrapper to use. If the metadata is embedded with the object it is not easily accessible. If the metadata is not embedded with the object, the metadata can be separated from the object when the item is transferred in some form. What would be ideal is a combination of the two, where the embedded metadata is automatically updated when the metadata in a database is changed. The two choices for wrapping AV metadata are METS and MPEG-21. METS, which grew out of the library field, is more widely adapted, while MPEG-21 was developed specifically with digital video and audio in mind, which might make it appealing to a more technically adept archive. Once again, any such choices and the way that the wrapper is employed need to be well documented.

The person who does that documentation leads to the other main way to ensure consistency and accuracy: who will be in charge of managing the metadata. Looking at their workforce, resources, and expectations for a digital collection, an organization needs to decide how the metadata will be collected and by whom. Metadata can be recorded at all steps along the item's lifecycle by a variety of departments (including users) or there can be dedicated cataloguers

[12] Wilson, et al, 83-89.

responsible for the activity. Similarly, an organization can hire outside consultants to set up the metadata structure and entry interface or they can keep someone on staff full-time to manage the metadata. For both of the former examples – a variety of enterers and a lower amount of oversight – there must be policies and techniques in place to automatically maintain consistency and accuracy of the metadata. These include training the employees who will be recording metadata, adopting controlled data dictionaries, having pull-down fields in the entry database, etc. However, even while working towards best practices, organizations that adopt such an approach will be missing out on the benefits of having what Priscilla Kaplan describes a data curator[13]. A data curator manages the metadata over time and can adjust the structure to technological or collection changes. Presence of a data curator means the metadata will be better prepared for future preservation needs and migrations into new metadata schemas.

Here, in conclusion, are some areas of further research for metadata and digital preservation. Understandably, digital preservation and the concomitantly required metadata is still an evolving process. Even in the more circumscribed world of moving image and sound collections the wide variety of format specific needs precludes a one size fits all solution. Still, the fact the it is becoming common practice to mix and match between metadata structures à la carte reveals the inadequacy of any one schema on its own to provide all of the descriptive, technical, and administrative metadata needs. Hopefully, while an imaginary MARC for digital multimedia metadata is both unlikely and not ideal, there will be greater standardization towards integrated metadata structures.

[13] Caplan, *Preservation Metadata*, 17.

Richard Gartner's *Metadata for Digital Libraries: State of the Art and Future Directions* is an important work as a first-step in that direction by pointing to schemas' shared structural base of XML[14]. Hopefully, as cultural institutions – here specifically addressing those with AV collections – grapple with this issue of structure combination and integration, they will not only document their work and findings for internal use, but make them public through registries and journal articles. More examples of real life case studies, and not just theoretical musings, will find shortcomings in mixing structures, aid others when they are conceiving metadata needs for digital collections, and hopefully constitute a sort of testing phase that will develop best practices for integration.

So, while an over-arching single metadata structure will not arise, what would be helpful would be more standardized integrated metadata schema related to the technical and preservation needs of subsets of the digital library/archive/preservation world. This would curb the constant experimentation and re-inventing of slightly different wheels that is currently going on. Perhaps something like Wilson et al's minimum requirement of metadata based on formats can be a starting point. One thing that would help in this would be adopting similar conceptual and semantic frameworks. For example, the five types of entity in PREMIS should not have to be mapped to the categories of OAIS information model – if they are the same concept they should be named the same. This next recommendation is perhaps a pipe dream, but it would be very useful for there to be a clearinghouse website that includes all of the studies

[14] Gartner, Robert, "Metadata for Digital Libraries: State of the Art and Future Directions," *JISC Standards and Technology Watch*, (April 2008), http://www.jisc.ac.uk/media/documents/techwatch/tsw_0801pdf.pdf.

on digital metadata structures that have been written. It would also include crosswalks. These suggestions imply the need not just for regulating standards for each individual schema but for inter-structure as well.

Finally, more automated means of harvesting metadata for digital AV content is needed to realize the potential of creating a good digital collection. The traditional cataloguing process in the analogue world is too time consuming for the deluge of digital AV content. For example, something analogous to optical character recognition, which has opened up the content of scanned text to very specific searches, is needed for video and audio. Clearly, such a system is at a much higher level of technical expertise. MPEG-7 mentions the possibility in its documentation and programs such as Autonomy Virage's VS Archive are already out there, but a more easily adoptable version should eventually be integrated into the digital preservation process.