DOI: 10.2352/issn.2168-3204.2016.1.0.161

Digital Preservation of Audiovisual-Based Materials: The State of the Art

Edward M. Corrado; University of Alabama Libraries; Tuscaloosa, Alabama, USA; Heather Moulaison Sandy; iSchool at the University of Missouri; Columbia, Missouri, USA

Abstract

In this paper, we explore the state-of-the-art for recommendations and best practices relating to long-term storage and preservation of audiovisual files. Without the proper support, audiovisual contents may become obsolete over time. Specifically, we examine file formats and metadata used in digital preservation of audiovisual resources. Additionally, we investigate the way these recommendations and best practices are integrated into three digital preservation systems.

Motivation

This paper examines the broad field of digital preservation as it pertains to audiovisual materials. Digital preservation is notoriously challenging, as it requires today's understanding of resources and technology to enable future access to digital objects. This challenge is especially felt when working with audiovisual objects; these objects are typically more complicated than textual documents or photographs and have fewer widely-held best practices for practitioners to implement.

The resources available to digital preservationists in this area are changing rapidly. This paper presents information on relevant file formats, metadata schema, and repositories that represent the state of the art in terms of the literature.

Problem

Digital preservation is ultimately a management problem, [1] with the preservation of audiovisual materials presenting additional challenges beyond these identified when working with more simple resources. This paper focuses on audiovisual materials, in particular audio files and video files, which are available in a variety of file formats; each of these formats involves certain advantages and disadvantages. Without the proper metadata, audiovisual resources are unfindable, and are unusable even if somehow discovered. Preservation systems are available to support the preservation work of information professionals, yet are changing all the time. What advantages and disadvantages can be identified for work with audiovisual materials for digital preservation when considering published best practices and recommendations?

Approach

Through a review of the published literature from 2014-2016 and an analysis of recent and current best practices for digital preservation relating to audiovisual materials, we present here a form of synthesis [2] of the state-of-the-art surrounding long-term access and use of audiovisual materials. We find that a number of resources are available, especially when adhering to recommended practices in terms of the file formats in which content is stored and its accompanying metadata. Below, we begin by presenting a review of current considerations and best practices relating to file formats, then to metadata, and finally, we give a snapshot of some systems of interest to digital preservationists working with audiovisual content.

Audiovisual File Formats for Digital Preservation

File formats are a standard, defined structure for how data is stored in a computer file. The structure of a file may include various content including a header, metadata, the content itself, and end-of-file indicators. [3] As a result, audiovisual file formats can be quite complicated. For example, MPEG-1 files may include encoded audio, video, and other data as well as methods to maintain synchronization between the different parts. Within the various parts, there can be subparts such as the color space, resolution and bitrate, frame/picture/block types, and more, that need to be defined for MPEG-1 video. With video, for example, if the images within a video file are using compression, it is not enough to have software that can play a particular file format, the software also needs to be able to decompress the video to play the specific file in question. Although each digital preservation scenario will produce unique requirements, by learning from each other and reviewing best practices, digital preservations can come up with an approach that works for their local institutions. Some examples of recommended practices follow.

The California Light and Sound Collection, an outgrowth of the California Preservation Program's California Audiovisual Preservation Project (CAVPP), works with both audio and video files. It created a list of technical specifications for the default output format of digitized recordings. In order to come up with this list, they reviewed current practice and solicited input for participating partner archives. [4] For digitized audio materials CAVPP prefers Broadcast WAV (24bit, 96 kHz. 2,304 Kbps for mono, 4,608 Kbps for stereo. L and R channels interleaved) for the masters and MP3 (160 Kbps for mono – 320 Kbps for stereo. L and R channels interleaved) for access copies. For digitized video they use Mov (10 bit uncompressed (4:2:2)) in a QuickTime wrapper for preservation and .mp4 (H.264 MPEG-4 Part 10) for access. It should be noted that for born digital audio and visual, they "maintain original specifications and embedded metadata." [5]

Like CAVPP, Biblioteca Europea di Informazione e Cultura (BEIC) works with sound files; it also recommends different file format specifications for different uses. Since BEIC mostly preserves commercially-produced CDs which they consider suitable for preservation and which they store in a safety vault, they do not typically produce a high-quality file. Adhering to international best practices, BEIC does not consider writable optical media such as CD-ROMS and DVD-ROMs as being suitable for preservation. For uses that require medium quality, including on-line listening, they use both MP3 and OGG files and will use lossless or lossy

compression. When using lossy compression they try not to diminish the audio quality more than necessary. For purposes such as preview listening where low-quality files suffice, they will use MP3 and OGG files with a high compression rate. [6]

Video files can also receive different treatment based on use-cases. Radiotelevisione Italiana (RAI) "manages both master quality broadcast formats (MXF) suitable for production, post-production or other re-use as well as lower quality proxy formats." [7] Since RAI is focused primarily on production, they do not concern themselves as much with preservation as some libraries, archives, and museums (LAMs), but they do believe that "as production and archive continues to converge, preservation will essentially become a production task." [7] BBC Scotland takes a similar approach in seeing its primary purpose as a producer and distributer of content, and it also follows industry standards for file formats. The Netherlands Institute for Sound and Vision's core "preservation format reflects its largest designated community, broadcasters; its master files are MXF and all material it manages is normalized to this format." [7] Institut National d'Audiovisuel (INA), which maintains French radio and television archives, chose JPEG200 in a MXF wrapper for a digital video preservation format. [7]

As demonstrated by some of the examples above, many LAMs prefer to have multiple file formats for each digital object to be used in varying methods instead of having one file format serve all purposes. While this approach requires LAMs to work with multiple versions of the same material, it is logical because long-term preservation and use are sometimes at odds. For example, downloading a high-quality digital preservation-worthy version of a video may not be practical on a mobile device but a smaller, compressed MPEG-4 (MP4) file will likely suffice in this instance.

Compression

CAVPP and others often prefer or recommend uncompressed audio or video. Kaur writes "Preservation formats should be free of any cryptography and compression techniques." [8] There are good reasons for this including the possibility of degradation of content and because compression algorithms may not be available in the future due to patent issues or other reasons. Another concern noted by CERN is that "A test with 10000 compressed files showed that with a likelihood of 99.8 % a SINGLE bit error makes the whole file unreadable, thus the data loss rate would be much higher for compressed files [compared to uncompressed files]." [9]

However, uncompressed files, especially video files, can be rather large. Although storage is becoming less expensive all the time, when preserving a large number of large files digital storage can still be expensive. Uncompressed files can also greatly increase network costs and upload times to cloud-based digital preservation systems compared to compressed versions. This is why some digital preservationists believe that, at least for large audiovisual files, the use of compression is necessary. One way to minimize the quality degradation in video archives while utilizing compression is to use lossless compression methods instead of being "influenced by the market-induced tendency to use lossy compression formats." [10] In order "To prevent loss of information, at least the relevant parts of the video must be stored in a high quality and should not be compressed with lossy algorithms." [11]

Collection Development/Accessioning

Most LAMs have a collection development or accessioning policy for physical collections. It is important to have similar policies in place to help determine what digital audiovisual materials should be preserved and for what length of time. Policy must reflect the mission of the LAM when determining technical requirements. For example, "the Royal Library of Sweden's mission includes maintaining comprehensive holdings of Swedish media history (television and radio programs) for reference purposes." [7] Since they are preserving these for reference purposes only, their "video collection in particular is of low technical quality." [7] Similarly France's Bibliotheque Nationale de France (BnF) has a reference only mission for video preservation, so they do not necessarily collect and preserve the highest quality video possible. [7]

These policies make sense because of the mission of the organizations involved and the purpose they are accessioning audiovisual materials. They also are less costly and, potentially, easier to deal with in the long term than other higher quality formats. This approach, however, is not without risk. It is impossible to know what future users of the materials may want to do with them. By not accessioning and preserving the highest quality versions, LAMs may inadvertently negatively impact or prevent future use.

Audiovisual File Format Polices

An analysis of digital preservation file format policies at Association of Research Libraries (ARL) member institutions published in 2014 included information about audiovisual file formats. The researchers found that digital preservationists are less likely to trust "file format types that do not have their roots in longstanding library digitization efforts." [12] However they did rank the top ten video formats they encountered in the policies, with Motion JPEG 2000, AVI, MPEG, MPEG-2, and Quicktime being the top 5. [12] For audio files, the top file formats in terms of relative confidence were Audio Interchange File Format, (AIF, AIFC, AIFF), Wave Audio File Format (WAV), Ogg Vorbis (OGG), MPeG-4 (Audio only; MP4, M4A), and MPEG-3 (MP3). [12, Appendix 2, p.34]

Emerging File Formats

New and emerging file formats and methods are always on the horizon. If possible, a good idea is to engage with researchers in your institution to see what file formats they are using. Another way is to look at digital preservation and digital asset management projects outside of your organization and to keep an eye on emerging trends in the marketplace.

Three dimensional (3D) still image and video is growing in popularity. In 2013 Autodesk and the Smithsonian Institute teamed up to create x3D Explorer, which "allows the Smithsonian to digitally preserve its extensive collection as interactive, 3D models." [13] This project also makes 3D files available to download in a variety of file formats. Some of them are in long-standing 3D file formats such as OBJ and STL. However there are also newly emerging 3D file formats that LAMs need to be aware of. Many of these are for use in 3D computer graphics applications. In many case, gaming applications are leading the way. One open format that is "designed to facilitate the transfer of complex scene data between applications such as modeling tools and game engines" [14] is the Open Game Engine Exchange (OpenGEX) format. Two other open formats in this arena are X3D, an XML-based file format which is designed to represent computer graphics in 3D and Alembic, an interchange format for computer graphics that has been widely adopted by visual effects and animation professionals.

Although not a file format, technologies such as IRENE (Image, Reconstruct, Erase Noise, Etc.) are also of interest for audio

DOI: 10.2352/issn.2168-3204.2016.1.0.161

preservation. IRENE is an "audio digitization technique […] developed by Dr. Carl Haber at the Lawrence Berkeley National Laboratory." [15] Instead of directly creating a digital audio file, IRENE creates a high-resolution digital images (either 2D or 3D depending on the application) of wax disks, lacquer transcription disks, record albums and other media. Software then is used to produce a digital audio waveform. In this case, assuming the software is available, the high resolutions images which may be in the TIFF file format, which is ubiquitous in digital preservation, can be used as the digital preservation master.

Metadata Supporting Digital Preservation

Metadata can be defined as structured information about a resource, and version 3.0 of the PREMIS Data Dictionary defines *preservation metadata* "as the information a repository uses to support the digital preservation process." [16, p.2] A variety of metadata, including descriptive metadata, administrative metadata, and preservation metadata, are all required for retrieval and long-term use of AV files. Given the complexities of metadata for audiovisual files, in light of the literature, this section will focus on best practices in metadata standards and schemas.

Descriptive Metadata

Descriptive metadata used in database environments "describes a resource for purposes such as discovery and identification"; [17] it identifies important elements such as the author, the title, and potentially the topic of a resource. *Controlled vocabularies* such as the ones maintained by the Getty (e.g. Union List of Artist Names; Art & Architecture Thesaurus) can be useful in providing consistent access to resources. One theme that continues to emerge when metadata is discussed, is that "There is no unique metadata standard sufficient to describe all the documents emerging in various kinds of formats, helpful in efficient information retrieval"; [18] thus, a variety of options and solutions are presented in the literature.

Encoding schema can be considered a kind of descriptive metadata in the sense that they provide a set of labels, or elements, to be used in metadata records. [19]

The most widely used encoding schema is the Dublin Core Metadata Initiative (DCMI). The strength of Dublin Core (DC) lies in its simplicity, so it is not only able to be used for audiovisual resources. Because it is intended for a wide variety of audiences, however, DC might not be the only schema required to describe audiovisual content.

At present, metadata records are often encoded in XML, using standards that are adapted to the community of users and to the kind of resource being described. Libraries and other cultural heritage institutions might use the Metadata Encoding and Transmission Standard (METS) (http://www.loc.gov/standards/mets/), an XML-compatible schema "encoding descriptive, administrative, and structural metadata regarding objects within a digital library" [20] to provide metadata for a variety of resources in their repositories; METS is currently in version 1.11 [20] and supports the AudioMD and VideoMD formats (which also function as METS or PREMIS extensions) from LC. [18]

A number of schema have developed to support video description that will be useful as well in preservation. PBCore 2.1 (Public Broadcasting Metadata Dictionary), for example, is adapted to work with sound and moving images. The United States' National Archives and Records Administration (NARA)'s AVI MetaEdit "inserts essential metadata into selected video files such as information about how the file was made." [21] Another tool by NARA, called reVTMD, provides "structure to organize important technical information, such as how big the image is or how fast it should be playing, in a reusable XML format." [21]

Audio files also require special considerations due to the complexity of the formats, especially since so much of their metadata may be embedded. For example, the Broadcast Wave Format (BWF), not unlike METS, has been around for a long time and is a kind of data carrier (in this case, specifically for audio) that can be embedded in .wav files [22] through the use of the open source BWF MetaEdit tool available in SourceForge.[23] For a number of years, information professionals have been awaiting the Audio Engineering Society's AES-X098 standard to support audio preservation, but its status is, at present, unclear in the literature.

Web-Based Description

Outside of the traditional repository environment, a few metadata schema of note are providing description for web-based content and will potentially be relevant to web-archiving efforts. Schema.org, the collaborative initiative spearheaded by Google, Bing, Yahoo!, and Yandex, supports microdata and the inclusion of description throughout web documents. [24] Additionally, social media can capitalize on web-based metadata. Facebook, for example, has developed its own Open Graph protocol (OGP) as a set of elements to describe resources, including audiovisual resources, on the open web. When web-based content is embedded in a Facebook status update, certain descriptive elements are imported via OGP. [19] OGP, however, can only be placed in a document's header, limiting its ability to self-describe. [24]

Administrative Metadata

Administrative metadata ensures that a resource can be used into the future. Pomerantz defines administrative metadata as metadata that "provides information about the origin and maintenance of an object: for example, a photograph might have been digitized using a specific type of scanner at a particular resolution, and might have some copyright restrictions associated with it." [19, p. 17]

Rights metadata helps LAMs and their users make use of content. Acknowledging the importance of balancing the needs of repositories and users, a working group composed of members from the Digital Public Library of America (DPLA) and Europeana identifies five characteristics of effective rights statements: "1) simple, 2) flexible, 3) descriptive, 4) accurate and 5) transparent." [25] Accordingly, their new International Rights Statements has been developed and was published in October 2015. This new recommendation complements two common rights schemas that have already been in existence for a number of years: RightsMD from the Library of Congress which works with LC's METS. Also, copyrightMD (http://www.cdlib.org/groups/rmg/) from the California Digital Library allows for rights metadata to be added to a record and can also be used with METS.

Provenance metadata "can be used for many purposes, such as understanding how data was collected so it can be meaningfully used, determining ownership and rights over an object, making judgements about information to determine whether to trust it, verifying that the process and steps used to obtain a result complies with given requirements, and reproducing how something was generated." [26] To do this effectively, agents and activities, among other things, are recorded. digiprovMD from LC can record provenance information, and can be used with METS. PROV is a W3C standard for recording provenance metadata.

Technical metadata supply needed information about the file. In her 2014 article on administrative metadata in libraries, Otto finds that only 35% of respondents who accept video reported that their systems accommodated video technical metadata. The most common metadata was 1) duration; 2) video encoding scheme; 3) codec information (name, version, creating app, etc.); 4) presence of sound; and 5) sampling information (sampling rate, bit depth, word size, etc.). Similarly, with audio files, only 38% of library repositories that accepted audio could accommodate audio technical metadata. The most common elements were 1) duration; 2) audio encoding (e.g., PCM); 3) audio codec information (name, version, creating app, etc.); 4) audio bit rate information (kBps, whether fixed or variable, etc.); and 5) audio sampling information (sampling rate, bit depth, word size, etc.)." [27]

Other metadata most closely aligned with the files themselves and with having the files render correctly in systems for users. *Structural metadata* "indicates how compound objects are put together" [17] and has not been a recent topic of scholarly discussion in the literature.

Preservation Metadata

Preservation metadata supports the long-term access and use of content. The primary preservation metadata schema is PREMIS (PREservation Metadata: Implementation Strategies), now in version 3.0. The newest major revision of the Data Dictionary was released in June 2015 and was last updated in November 2015 (http://www.loc.gov/standards/premis/v3/index.html).

The PREMIS 3.0 data model identifies four entities: [Digital] Object (or an Environment supporting a Digital Object), Rights Statement, Agent, and Event. In the PREMIS model, the framework is implemented as semantic units under each of the identified entities.

To the Object entity, a preservationLevelType semantic unit has been added in PREMIS 3.0. This allows digital preservationists to specify the level of commitment on the part of the institution for a number of preservation activities that will be relevant to audiovisual materials. For example, an institution may have a different level of commitment for bit preservation than for logical preservation. [16], [28]

To the Agent entity in PREMIS 3.0, a new semantic unit was added called agentVersion to allow the version of software to be recorded. [16], [28]

Finally, to the Event entity, eventDetailInformation was added in PREMIS 3.0 as a semantic unit to provide more information about an Event. eventDetail is now available in the container eventDetailInformation along with an extension for eventDetailExtension to allow for more granularity. [16], [28]

Digital Preservation Systems Overview

File formats and metadata can concretely be explored in relation to selected digital preservation systems. Below we briefly investigate three such systems.

DuraSpace's DuraCloud

DuraSpace's DuraCloud (http://www.duracloud.org/) is an open source platform that uses the cloud for storage [29] and that offers a hosted version that systematically uses two cloud vendors. [30] Current features include audio and video streaming; it is "based on Amazon's Cloudfront service which uses the Flash Media Server to host streaming files over RTMP in MP3, MP4, FLV, and other formats" [31] In its white paper from 2014, the POWRR grid project (funded by the Institute of Museum and Library Services (IMLS)) reported one reviewer of the DuraCloud project noted the attractiveness of this streaming feature. Another reviewer, however, noted ""very minimal metadata ... (properties and tags)" capacity produced "minimal functionality" for those seeking to add and manage metadata for digital objects. [32]

Archivematica

Archivematica (https://www.archivematica.org/en/), an open source preservation system conforming to OAIS, permits the use of a number of the standards described and actively seeks funding to adapt other standards. [28] Format policies indicate that "preservation formats must all be open standards." [33] Archivematica is currently in 1.4. "Archivematica is compatible with METS, PREMIS, Dublin Core and other best-practice metadata standards." [32] For example, Archivematica's website demonstrates that it permits PREMIS and METS to be used with common vocabularies. For example, Archivematica requires techMD as a "PREMIS:OBJECT" and digiprovMD as a "PREMIS:EVENT", thereby combining standards described above. [35] Upfront planning requires institutions to configure their Format Policy Registry (FPR), requiring them to give a lot of thought to normalization, which happens on ingest, and future file format needs. This observation is consistent with Sprout and Jordan, who note that some institutions lost considerable time in getting their collections ingested simply due to the fact that the institutions did not have comprehensive preservation policies and they needed to be written first. [36]

Libnova's Libsafe

Libnova (http://www.libnova.com/en/) is a company that provides digital preservation solutions to the cultural heritage sector. The company started in Spain in 2009 and expanded internationally in 2014. One of Libnova's digital preservations solutions is Libsafe, an OAIS and ISO 14,721 compliant digital preservation system that can be hosted in the cloud (or locally). It is designed to make "difficult tasks easy." [37] Libsafe is a flexible system that can preserve any file format, identifying and validating over 1,400 of them, and allows for both standard and custom metadata schema. [38] Although custom metadata schemas are possible, Libsafe relies on METS protocols and another standards for consultations and extractions. [39]

Observations

Although these and other digital preservation systems may accept varying forms of metadata schemas, extensive customizations may be required in order to take advantage of them. And while they can accept different file formats, the services and support for validation, normalization, and other activities may be limited when using file formats that are not as widely used for preservation. Another issue to consider when using hosted digital preservation systems is the amount of Internet bandwidth required to transfer large audiovisual files across the Internet to the system.

Discussion and Limitations

The literature and best practices that informed this paper were relevant and sufficient. Although more would have been better, there is enough published for our analysis to provide a snapshot of current and emerging best practices. Although digital preservationists of audiovisual materials are working through these very real and very technical problems every day, relatively little appears in the scholarly literature to guide best practices. Digital preservationists may not be developing the standards, but they are

key stakeholders in the way file formats, for example, support use and long-term use. Large institutions are doing their part by publishing case studies, for example, detailing their use of formats; and some librarians are researching and documenting the use of metadata. More needs to be done in this rapidly-changing field, however, to inform practice in a large variety of institutions through the consistent publication of articles, white papers, books, etc., by those involved in all aspects of the work, and at all sizes of institutions.

Conclusion

In this paper, we examined the current literature pertaining to file formats and metadata supporting digital preservation of audiovisual materials. In demystifying the variety of available resources on this challenging topic, this paper gives insight into what is available, what might be coming, and how this might play out in individual LAM digital preservation initiatives.

References

[1] E. M. Corrado and H. L. Moulaison, Digital Preservation for Libraries, Archives, and Museums, Lanham: Scarecrow Press, 2013.

[2] H. Cooper, L. V. Hedges, and J. C. Valentine (Eds.), The Handbook of Research Synthesis and Meta-Analysis. 2nd edition, New York: Russell Sage Foundation, 2009.

[3] P. Christensson. "File Format Definition." [Online]. Available: http://techterms.com/definition/file_format, Mar. 15, 2011.

[4] R. P. Hulser. (2015, April). "The California Light and Sound Collection: Preserving our Media Heritage." *Computers in Libraries* vol. 35, no. 3. pp. 4-10.

[5] "Target Audio and Video Specifications: California Audiovisual Preservation Project (CAVPP) For Quality Assurance." Internet: http://calpreservation.org/wp-content/uploads/2014/12/CAVPP-file-specs-2014.11.20.pdf, Nov. 26, 2014. p. 1.

[6] A. Baratè, L. A. Ludovico, and G. Haus, "Integration of Audio Resources into a Digital Library: The BEIC Case Study," *The International Journal of Digital Curation*, 2015, vol. 10, no. 2, pp. 48-57. doi:10.2218/ijdc.v10i2.345

[7] B. Delaney and A. de Jong, "Media Archives and Digital Preservation: Overcoming Cultural Barriers." *New Review of Information Networking*, vol. 20, no. 1/2 (2015): 73-89. doi:10.1080/13614576.2015.1112626.

[8] H. Kaur, "Digital preservation of manuscripts: An Indian perspective with special reference to Punjab," in *Emerging Trends and Technologies in Libraries and Information Services (ETTLIS), 2015 4th International Symposium on* (Jan. 6-8, 2015): 271-274. doi:10.1109/ETTLIS.2015.7048210.

[9] B. Panzer-Steindel. "Data Integrity: Draft 1.3." Internet: https://indico.cern.ch/event/13797/session/0/contribution/3/attachments/115080/163419/Data_integrity_v3.pdf, Apr. 8, 2007. p. 3.

[10] N. Orio, L. Snidaro, S. Canazza, and G. L. Foresti, G. L. (2009, Dec.). "Methodologies and tools for audio digital archives." *International Journal On Digital Libraries*, vol. 10, no. 4 (Dec. 2009): 201-220. doi:10.1007/s00799-010-0060-6.

[11] A. Uherek, S. Maier, and U. M. Borghof. "Requirements for Long-Term Preservation of Digital Videos and First Experiments with an XMT-Based Approach" in EUROCAST 2015, LNCS 9520, R. Moreno-D´ıaz et al. (Eds.): pp. 203–210, 2015. doi:10.1007/978-3-319-27340-2 26.

[12] K. Rimkus, T. Padilla, T. Popp and G. Martin, "Digital Preservation File Format Policies of ARL Member Libraries: An Analysis." *D-Lib Magazine,* vol. 20, no. 3/4 (2014, March/April). http://www.dlib.org/dlib/march14/rimkus/03rimkus.html.

[13] V. Quirk. "Autodesk Launches New Tool for Digital Preservation." *ArchDaily*. [Online]. (Nov. 29, 2013) Available: http://www.archdaily.com/452845/autodesk-launches-new-tool-for-digital-preservation. [Accessed: Jan. 30, 2016].

[14] E. Kengyel. "Open Game Engine Exchange (OpenGEX)." [Online]. Available: http://opengex.org/. [Accessed: Jan. 30, 2016].

[15] S. Norris. (2014) "Toward an Ontology of Audio Preservation." Journal of the American Institute for Conservation, 53(3), 171-181. doi:10.1179/1945233014Y.0000000026 (page 178).

[16] PREMIS Editorial Committee, PREMIS Data Dictionary for Preservation Metadata: Version 3.0. 2015. http://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf.

[17] National Information Standards Organization (U.S.), Understanding Metadata. Bethesda MD: NISO Press, 2004. http://www.niso.org/publications/press/UnderstandingMetadata.pdf. p. 1.

[18] P. Ramesh, J. Vivekavardhan, and K. Bharathi, "Metadata Diversity, Interoperability and Resource Discovery Issues and Challenges," DESIDOC Jour. of Library & Information Technology, vol. 35, no. 3, pp. 193–199, Apr. 2015. p. 198.

[19] J. Pomerantz. Metadata, Cambridge MA: MIT Press, 2015.

[20] "Metadata Encoding and Transmission Standard (METS) Official Web Site." [Online]. Available: http://www.loc.gov/standards/mets/. [Accessed: 30-Jan-2016]. Intro.

[21] NARA. "Sharing the Video Metadata Love." NARAtions, February 28, 2012. [Online]. Available: http://narations.blogs.archives.gov/2012/02/28/sharing-the-video-metadata-love/. para. 5.

[22] D. O. Melvin, "Managing Metadata Interoperability within Audio Preservation Framework: Integrating the Metadata Encoding & Transmission Standard (METS) and Multichannel Source Material into Digital Library Audio Collections," *Library Philosophy and Practice* (e-journal). May 2014. Paper 1117.

[23] "Guidelines: Embedding Metadata in Broadcast WAVE Files - Federal Agencies Digitization Guidelines Initiative." [Online]. Available: http://www.digitizationguidelines.gov/guidelines/digitize-embedding.html#tool. [Accessed: Jan. 30, 2016].

[24] D. Scott, "White Hat Search Engine Optimization (SEO): Structured Web Data for Libraries," *Partnership: The Canadian Journal of Library and Information Practice and Research*, vol. 10, no. 1, 2015.

[25] International Rights Statements Working Group, "Requirements for the Technical Infrastructure for Standardized International Rights Statements." Oct-2015. p. 12.

[26] "PROV Model Primer." 2013. [Online]. Available: https://www.w3.org/TR/2013/NOTE-prov-primer-20130430/. [Accessed: Jan. 30, 2016]. 1. Introduction, para. 1.

[27] J. J. Otto, "Administrative Metadata for Long-Term Preservation and Management of Resources," *Library Resources & Technical Services*, vol. 58, no. 1, pp. 4–32, 2014.

[28] A. Dappert, "Digital Preservation Metadata and Improvements to PREMIS in Version 3.0," presented at the DCMI/ASIST Joint Webinar, May 27, 2015. http://www.loc.gov/standards/premis/v3/tutorial.html.

[29] M. Kimpton and C. M. Morris. "Managing and Archiving Research Data: Local Repository and Cloud-Based Practices." In *Res. Data Management: Practical Strategies for Information Professionals*, 2014.

[30] Y. Han. "Cloud Storage for Digital Preservation: Optimal Uses of Amazon S3 and Glacier." *Library Hi Tech* vol. 33, no. 2 (June 15, 2015): 261–71. doi:10.1108/LHT-12-2014-0118.

[31] "Video and Audio Streaming | DuraCloud." [Online]. Available: http://www.duracloud.org/video_and_audio_streaming. [Accessed: Jan. 28, 2016]. para. 1.

[32] J. Schumacher, L. M. Thomas, D. VandeCreek, S. Erdman, J. Hancks, A. Haykal, M. Miner, P.-A. Prud'homme, and D. Spalenka. "From Theory to Action: 'Good Enough' Digital Preservation Solutions for Under-Resourced Cultural Heritage Institutions: A Digital POWRR White Paper for the Institute of Museum and Library Services," August 2014. http://commons.lib.niu.edu/handle/10843/13610. p. 8.

[33] "Technical Architecture | Documentation (Archivematica 1.5 (beta)) | Archivematica: Open-Source Digital Preservation System." [Online]. Available: https://www.archivematica.org/en/docs/archivematica-1.5/user-manual/overview/technical/#technical-arch. [Accessed: Jan. 28, 2016].

[34] J. Schumacher, L. M. Thomas, D. VandeCreek, S. Erdman, J. Hancks, A. Haykal, M. Miner, P.-A. Prud'homme, and D. Spalenka. "From Theory to Action: 'Good Enough' Digital Preservation Solutions for Under-Resourced Cultural Heritage Institutions: A Digital POWRR White Paper for the Institute of Museum and Library Services," August 2014. http://commons.lib.niu.edu/handle/10843/13610. p. 8.

[35] "METS Schematron Rules - Archivematica." [Online]. Available: https://wiki.archivematica.org/METS_schematron_rules. [Accessed: Jan. 28, 2016].

[36] B. Sprout and M. Jordan. "Archivematica As a Service: COPPUL's Shared Digital Preservation Platform." *Canadian Journal of Information & Library Sciences*, vol. 39, no. 2 (2015), 235-244.

[37] "About Us. Libnova. Technology Changes, Information Prevails." [Online]. Available: http://www.libnova.com/en/about-us/. [Accessed: Feb. 1, 2016].

[38] "Libsafe Digital Preservation Software." [Online]. Available: http://www.digitalpreservationsoftware.com/digital-preservation-solutions/libsafe-digital-preservation-software/. [Accessed: Jan. 31, 2016].

[39] "Libsafe." [Online]. Available: http://www.digitalpreservationsoftware.com/wp-content/uploads/documents/20140201-libsafe-catalog-eng.pdf. [Accessed: Jan. 31, 2016].

Author Biography

Edward M. Corrado is Associate Dean, Library Technology Planning and Policy at the University of Alabama where he is responsible for information technology operations in the University Libraries as well as planning and executing a strategic vision for library technology. He has published and presented nationally and internationally on various topics including digital preservation, cloud computing, and open source software in libraries. Mr. Corrado received his Masters of Library Services from Rutgers University.

Heather Moulaison Sandy is assistant professor at the iSchool at the University of Missouri. She has published twenty peer-reviewed articles and has presented extensively nationally and international on topics relating to digital technologies and organization of information. Dr. Moulaison Sandy received her PhD in Library and Information Science at Rutgers University, and received a Masters of Arts in French and a Masters of Science in Library and Information Science from the University of Illinois at Urbana-Champaign.

