

Project Title: Digitization for Everybody (Dig4E): Modular Educational Resources for Interpreting Standards for Still Image and Audio-visual Digitization

Proposed Activities and Rationale for the Project

Motivation and need: The primary motivation for the proposed project by the University of Michigan School of Information originates from insights gained over a decade of researching the long arc of digitization practices in libraries and archives and teaching graduate level courses on digital imaging and the preservation of audio-visual resources.¹ Research in the past decade by Mats Dahlstrom, Paul Conway, Zinaida Manzuch, and others has highlighted the importance of human agency in establishing the technical parameters of the digital products created through digital conversion.² Teaching at UMSI by PI Paul Conway has exposed a major gap between the technical complexity, sophistication, and maturity of many of the international standards for digitization and the knowledge required to be a professional practitioner in the cultural heritage sector. In an effort to fill this gap with existing technical and research literature, we have reached out to leading standards developers associated with the Federal Agencies Digital Guidelines Initiative (FADGI), the International Association of Sound and Audiovisual Archives (IASA), the Library of Congress, the British Library, and elsewhere. These experts (some of whom have agreed generously to contribute their time and talent to the proposed project) share an interest in closing this knowledge gap for new and mid-level information professionals, beyond the range of random courses and occasional workshops. We need an approach to knowledge transfer that promises a wide reach, communicates in clear language with ample examples, and includes self-paced and flexible learning with options for self-assessment.

The fundamental assumption of the project is that the underlying international standards for still image and audio and video digital transformations (including metadata) are largely in place and quite stable (still image and analog audio) or clearly settling into place (analog video). As promulgated and maintained by a variety of US and international standards bodies, these technical standards are generally difficult to understand, use, and interpret in the classroom and in practical local contexts. One consequence, we find, is that existing and emerging standards are not applied as well or as consistently as they could or should be. Technical documentation is rich, while supporting explanatory material is weak or non-existent. The complexity of existing standards and guidelines for digitization is a major conundrum for long-term preservation, because the justification for digital preservation turns in part on the value of the content to be preserved. Standards-based digitization is one of the principal mechanisms for creating content worth preserving.³

¹ Paul Conway, "Best Practices for Digitizing Photographs: A Network Analysis of Influences." *Proceedings of IS&T Archiving 2008*, Imaging Science & Technology, Berne, Switzerland, June 24-27, 2008, pp. 94-102.

² Mats Dahlstrom, "Critical Editing and Critical Digitization," in *Text Comparison and Digital Creativity*, edited by Wido Th. Peursen, Ernst Thoutenhoofd, and Adriaan Weel. Brill Online, 2010; Paul Conway, "Building Meaning in Digitized Photographs." *Journal of the Chicago Colloquium on Digital Humanities and Computer Science* 1 (no. 1) 2009; Zinaida Manzuch, "Ethical Issues in Digitization of Cultural Heritage," *Journal of Contemporary Archival Studies* 4 (2) 2017. <http://elischolar.library.yale.edu/jcas/vol4/iss2/no4>.

³ Paul Conway, "Preservation in the Age of Google: Digitization, Digital Preservation, and Dilemmas." *The Library Quarterly* 80 (1) 2010: 61-79.

We cannot predict the long-term technological future with any certainty, but we know that libraries, archives, and museums are living on borrowed time when it comes to the rescue and digital reformatting of analog audio-visual resources. Although technical standards for digital reformatting of audio-visual resources are sufficiently stable to promulgate and teach them, in practice the existential threat to audio-visual heritage is driven by a combination of its dependency on sometimes-obsolete playback equipment, the fragility of storage media created prior to about 1980, and the time it takes to read, assess, and reformat time-based media. Cultural heritage professionals must act within the framework of existing standards, make difficult appraisal decisions, and dedicate limited resources to saving what can be saved.

In the area of still image digitization, we also know that services to end users increasingly turn on the existence and availability of digital image surrogates of cultural heritage resources.⁴ Research by Paul Conway and Ricardo Punzalan exposed continuing debates in the archival and library professions over the gains and losses in the process of digital imaging. From the perspectives of users, they found that scholars, professional researchers, and expert avocational researchers have overwhelmingly strong preferences for the affordances that the digital images and digital data bring to archival collections.⁵ Users of digitized cultural heritage resources online, especially users with domain expertise on their subject of inquiry, are increasingly demanding access to standards-conforming archival surrogates.

For purposes of this project, the preservation and digital reformatting of motion picture films and the accompanying sound sources is excluded. Motion picture digitization currently is fraught by the existence of many different and competing standards, with little clarity yet on mechanisms for reconciling the tremendous variability of historical film (image and sound) with emergent standards for digital cinema projection. According to film scholar Giovanna Fossati, the “turmoil” around the existing logics of production, distribution and exhibition of films “has spread from the film industry to its audiences, from academia to cultural institutions.”⁶ The FADGI Audio-Visual Working Group recognizes that the state of standards development regarding the digitization of motion picture film is far from settled, largely due to the wide array of original source materials, disagreement on possible digital output formats, and ongoing vigorous debate on long-term preservation on film versus digital files.⁷

The three fundamental barriers that librarians, archivists, and museum professionals face in providing such high quality digital surrogates of their paper-based and audio-visual collections are: 1) time, 2) money, and 3) technical knowledge. Time is the enemy of fragile and obsolete media; there will never be enough money to rescue, extend the lifespan, and/or digitally reformat

⁴ Paul Conway, “Digital Transformations and the Archival Nature of Surrogates,” *Archival Science* 15 (1) 2015: 51-69. DOI: 10.1007/s10502-014-9219-z

⁵ Paul Conway and Ricardo Punzalan, “Fields of Vision: Toward a New Theory of Visual Literacy for Digitized Archival Photographs.” *Archivaria* 71 (Spring) 2011: 63-97. Winner: Hugh A. Taylor Prize, Association of Canadian Archivists.

⁶ Fossati, Giovanna, *From Grain to Pixel: The Archival Life of Film in Transition*. Amsterdam: Amsterdam University Press, 2011, p. 15.

⁷ Federal Agencies Digital Guidelines Initiative, Audio-Visual Working Group. “Digitizing Motion Picture Film: Exploration of the Issues and Sample SOW,” April 18, 2016.

http://www.digitizationguidelines.gov/guidelines/Motion_pic_film_scan.html

everything of long-term value. This proposal, therefore, addresses a basic need of nearly the entire cultural heritage community for knowledge on how to master, adapt, and apply existing and emergent technical standards. The primary need is to reach and convince a wide audience to cross a learning bridge between the technical complexities of international digital conversion standards and local adaptive practices that require higher levels of knowledge and skill than are present in most cultural heritage organizations, even after two decades of efforts to develop best practice guidelines. Part of the challenge addressed in this proposal is advocacy for embracing the technical qualities embedded in relevant preservation-oriented standards. The other part of the challenge addressed in the proposed project is making technical standards intellectually accessible and understandable to professionals who are not themselves technical experts. The project seeks to develop in students and professionals a form of conceptual and technical literacy that enables them to interpret existing and emergent standards and assess their application in practice.⁸

Standards domains for the project: The three media domains under consideration in this proposal are still image digitization of visual resources, analog audio resources conversion, and analog videotape conversion. In these three domains, supporting technologies and the underlying national and international standards either have stabilized or are on the verge of being sufficiently stable to inspire confidence in the long-term viability and usefulness of the digital files produced in the conversion process. The level of standards maturity varies by medium:

Still image digitization is the most highly standardized, stable, and mature conversion process, as evidenced by a standards framework developed almost a decade ago that is still highly relevant today.⁹ Yet the underlying standards for specifying and measuring the resolution, tonal values, and color fidelity of image files (e.g., ISO International Standards Organization (ISO) 12233 Spatial Frequency Response; ISO 14545 Opto-Electrical Conversion Function) are the least well understood in practice or teachable in the classroom. Current guidelines developed by the Federal Agencies Digital Guidelines Initiative (FADGI) are dense and challenging to adopt, in part because they represent an admirable effort to embed knowledge of image science metrics in workflows previously dependent upon subjective visual inspection.¹⁰ Applying these guidelines in practice requires making the connections between the characteristics of original source materials, the technical performance of the entire digitization workflow, and the purposes to be made of the created in the digitization process.¹¹ We have partly built the proposed project around helping practitioners make these connections.

Audio digitization guidelines developed by the International Association of Sound and

⁸ Don Williams and Peter D. Burns, "Preparing for the Image Literate Decade" *Proceedings of IS&T Archiving 2009*, Society for Imaging Science and Technology, May 4-7, 2009, pp. 124-127.

⁹ FADGI. Still Image Group. Guidelines: Digital Imaging Framework, April 2, 2009. <http://www.digitizationguidelines.gov/guidelines/digitize-framework.html>

¹⁰ Federal Agencies Digital Guidelines Initiative. *Technical Guidelines for Digitizing Cultural Heritage Materials*, September 2016. <http://www.digitizationguidelines.gov/guidelines/digitize-technical.html>

¹¹ Paul Conway, "Overview: Rationale for Digitization and Preservation." *In Handbook for Digital Conversion Projects: A Management Tool*. Edited by Maxine Sitts. Andover, MA: Northeast Document Conservation Center, 2000. <http://www.nedcc.org/digital/ii.htm> Reprinted in: *Museums in a Digital Age*, edited by Ross Parry. London: Routledge, 2009.

Audiovisual Archives (IASA-TC 04) are more accessible (readable) to a non-specialized audience but are not tied particularly well to the rich array of international standards that anchor the guidelines.¹² Additionally, the absence of a glossary of terms (in the IASA documentation) and ties to the underlying technical literature make the IASA guidelines somewhat isolated and less convincing in practice, particularly in a way that avoids a “one size fits all” approach to audio digitization. The *Sound Directions* work at Indiana University is a partial but still incomplete bridge to international standards promulgated by the European Broadcast Union and the Audio Engineering Society, among others.¹³ The educational materials developed in the proposed project will complement and enhance the IASA guidelines while also focusing on classroom and workshop applications.

Videotape digitization is on the cusp of stabilization around agreed upon international standards for digital conversion, file formats, and the technical wrappers that facilitate exchange and access. IASA has released for comment a proposed synthesis of recommendations on the preservation of video recordings (IASA-TC 06), with the initial emphasis on the digital conversion of analog videotapes.¹⁴ The underlying standards, including the Material Exchange Format (MXF) for digital content and metadata wrappers,¹⁵ are technically complex and lack interpretive and explanatory materials that would make these standards accessible to students and professional practitioners in libraries and archives. The proposed project is an opportunity to get out ahead of these new video-oriented guidelines with explanatory and educational materials that could help foster the wider acceptance of video standards.

Other areas for project development

Technical metadata requirements are a common element in digitization standards documentation across all target media (visual resources, analog audio, and analog video). Technical metadata are those elements of metadata required to describe the technical characteristics of the digital file that results from digitization activity. Good technical metadata is required for present day interoperability and a fundamental component of long-term preservation. Such metadata is embedded in digital file formats, comprised of components of digital file wrappers, or structured in database driven XML schemes or in semantic frameworks for the web. Guidelines produced and distributed by FADGI and IASA point to the integral importance of technical metadata in the production of digital files from analog resources.

Existing guideline documents from both organizations offer little or no contextual information about technical metadata schemes and generally avoid detailed but necessary explorations of

¹² International Association of Sound and Audiovisual Archives. *IASA-TC 04 Guidelines on the Production and Preservation of Digital Audio Objects* (Second Edition, 2009). <https://www.iasa-web.org/tc04/audio-preservation>

¹³ Indiana University. *Sound Directions: Digital Preservation and Access for Global Audio Heritage* (2007-08). <http://www.dlib.indiana.edu/projects/sounddirections/>

¹⁴ International Association of Sound and Audiovisual Archives. *IASA-TC 06 Guidelines for the Preservation of Video Recordings* (forthcoming 2018 and 2019). <https://www.iasa-web.org/tc06/guidelines-preservation-video-recordings>

¹⁵ FADGI. *MXF Application Specification*. http://www.digitizationguidelines.gov/guidelines/MXF_app_spec.html ; FADGI. *Digital File Formats for Videotape*. http://www.digitizationguidelines.gov/guidelines/video_reformatting_compare.html

the relevant metadata standards for digital imaging and audio-visual conversion. The level and usefulness of explanatory and interpretive guidance is inadequate for most learning and training purposes. The proposed project includes the development of learning materials related to applicable technical metadata for digital still images, digitized analog sound recordings, and digitized analog videotapes. For digital still images, the relevant technical metadata standards are defined by ANSI/NISO Z39.87-2006 and the associated XML schemas maintained by the Library of Congress.¹⁶ For analog sound recordings, the Audio Engineering Society maintains the relevant standard for audio technical metadata as AES57-2011 (r2017).¹⁷ For digitized videotape, the Advanced Media Workflow Association, in close collaboration with FADGI has developed a proposed standard (AS-07) for complex multi-part video files that includes a core technical metadata set.¹⁸ It is likely that AS-07 will become the foundation for preservation-quality digitized video sources.

Other important components of preservation-quality digitization are the digitization workflow process itself, including a mastery of the tools for calibrating equipment, measuring system performance, diagnosing problems, and assessing the completeness and accuracy of technical metadata, particularly metadata embedded in digital files. People who wish to adopt best practices for digitization also need explanatory and support materials for tools used to read and diagnose embedded metadata in image, audio, and video files.

The project will focus on explaining / demonstrating the operations and outputs of software tools for image quality and metadata completeness. Particular attention will be devoted to conveying skills in using and interpreting tools developed by AVPreserve and other groups in collaboration with FADGI (Open DICE [Digital Imaging Conformance Evaluation]),¹⁹ the National Archives and Records Administration (AVI MetaEdit for metadata management), and the Library of Congress (BWF MetaEdit for inspecting embedded metadata in the Broadcast Wave Format). The Open DICE imaging evaluation toolkit is itself an open source derivative of the mostly proprietary tools (GoldenThread, GoldenTouch) developed by Image Science Associates, under the leadership of image scientist Don Williams.²⁰ AVPreserve's tools for assessing risk and value are promising new developments, including MediaRivers for assessing the relative intellectual value of collections of audio-visual resources and MediaScore for generating a measure of risk to these collections.²¹ Neither tool comes equipped with explanatory documentation optimized for teaching or self-instruction.

Audiences for Dig4E: The audiences and primary beneficiaries for Dig4E (Digitization for Everybody) are teachers and learners, broadly defined, who wish to understand and then apply standards for digital still images, audio and video conversion, and the associated core technical

¹⁶ Library of Congress. NISO Metadata for Images in XML Schema (MIX). <https://www.loc.gov/standards/mix/>

¹⁷ Audio Engineering Society. AES57-2011 (r2017): AES standard for audio metadata – Audio object structures for preservation and restoration. <http://www.aes.org/publications/standards/search.cfm?docID=84>

¹⁸ FADGI. AS-07: MXF Archive and Preservation Format Application Specification, September 8, 2017. http://www.digitizationguidelines.gov/guidelines/MXF_app_spec.html

¹⁹ FADGI. Open DICE. <http://www.digitizationguidelines.gov/guidelines/digitize-OpenDice.html>

²⁰ Image Science Associates. <http://www.imagescienceassociates.com>

²¹ AVPreserve. Products. <https://www.weareavp.com/products/>; Indiana University. *MediaRivers and MediaScore*. <https://www.weareavp.com/indiana-university-announces-release-of-mediascore-and-mediarrivers/>

metadata elements.

Teachers: Digitization is taught at the graduate level in iSchools as stand-alone courses (e.g., Michigan, Syracuse, Texas), as modules in courses on collection development (Washington, Pratt, Toronto), digital curation (Maryland, UNC Chapel Hill, Toronto), and library/archives technology (Illinois, Simmons, San Jose State University). Instructors in graduate level courses do not have the benefit of teaching resources explicitly developed for classroom teaching or assessable assignments in these classes. We have designed the proposed project explicitly to generate and test educational resources for graduate level instruction.

Workshop Instructors: Workshops through Lyrasis, the Northeast Document Conservation Center, as well as preconference offerings at professional associations draw on standards work, but largely convey best practices at a generalizable level, rather in a way that encourages and supports the adaptation of guidelines for local settings. The proposed project will develop and test educational materials for learners in workshops or the workplace, along with self-assessment components that learners can use to judge their progress through the materials.

Managers and Trainers in the United States: The resurgence of in-house digitization labs in libraries and archives (due, in part to declining costs of equipment and the stabilization of conversion standards) increasingly depends on seasonal student labor to “turn the pages” or “playback the media.” The training of temporary workers is a major and costly challenge for in-house digitization labs. We have designed this project to generate educational materials on underlying digital imaging standards and guidelines that can be adapted for training students and temporary workers in local contexts. For digitization projects that involve outsourcing of work to third party providers of audio-visual conversion services, the project will provide a grounding in digitization standards that learners can work into requests for proposals (RFP). We will consult with some of those providers to advance the development of model RFP documents.

Independent Learners: Community archives and archivists are embracing digitization and audio-visual transformations, largely to support sharing and wider online access. However, based on a decade of involvement by the PI with community archives initiatives, most community-based programs focus their energies on the content and usefulness of the archival sources, often at the expense of technical matters, including digital imaging and metadata schemes that support exchange, interoperability, and long-term preservation. A wide understanding of digitization and metadata standards will provide knowledge and incentives that community collaboratives need to adopt the quality standards and guidelines taught in the project’s modules and reinforced through self-assessment.

Related Resources and Initiatives

The closest example of online resources on digitization best practices is a three-part series of lectures on “FADGI Metrics” presented by imaging scientist Don Williams as part of a fee-

based continuing education webinar sponsored by Lyrisis.²² These lectures, limited to still image digitization, are not freely available to a wide audience. In the UK, the recently launched #*dariaTeach* learning platform is an excellent model for delivering open source teaching materials. To date, this digital humanities initiative has not yet extended to instruction on standards-based digitization, although creative partnerships with this organization may well be an appropriate outcome of the proposed project.²³

To the best of our knowledge, no textbooks (printed or online) exist to support instruction or training in standards based digitization for any of the targeted media. In contrast to support for digital curation and digital preservation education, there are no “How-To-Do-It” manuals for standards-based digitization; and even if there were such published volumes, they would lack the sort of self-assessment of learning that we envision for the individual modules of this project. In the area of still image digitization, the combined proceedings of the annual IS&T Archiving conferences provide good guidance on emergent best practices but are too sprawling and unfocused for use in education and training settings.²⁴ Individual articles in the proceedings volumes will be useful additions to the proposed online resources. In the area of audio-visual archives a single volume by Anthony Cocciolo on moving image and sound collections, is an excellent general-purpose introduction to the issues of managing audio-visual collections in archival settings.²⁵ The volume contains a good glossary of terms, but does not provide a guide to the relevant digital conversion standards.

The standards generating communities of interest, particularly FADGI and IASA, provide a variety of useful technical resources that do not themselves comprise a collection of ready-to-teach materials for students. The website for FADGI contains important documentation on the evolution of still image and audio-visual digitization, along with a very useful glossary of terms relating to digitization across a variety of media.²⁶ There is no similar glossary on the IASA website, although this organization makes its guidelines publications freely available through the web. Individual guidelines documents define terms and may occasionally provide a limited glossary.

Far more problematical is the lack of open access to standards documents maintained and published by standards organizations themselves. With few exceptions, standards documents from the International Standards Organization, the European Broadcast Union, and the Audio Engineering Society, to name a few of the relevant standards bodies, are held behind a steep paywall. One major purpose of the project is to highlight appropriate international standards while explaining how to read, understand, and interpret these standards. It is important to note that most of the underlying technology standards are not constantly in flux. The most important standards for still imaging are quite stable as are the underlying standards for the digitization of audio resources, particularly magnetic tape and common grooved media (45s

²² LYRISIS Course Catalog – F. <https://www.lyrisis.org/content/Pages/Courses.aspx?type=Letter&val=F>

²³ #*dariaTeach* [website]. <https://www.dariah.eu/teach/>

²⁴ Society for Imaging Science & Technology, Proceedings of Archiving Conferences. <http://ist.publisher.ingentaconnect.com/content/ist/ac>

²⁵ Anthony Cocciolo, *Moving Image and Sound Collections for Archivists*. Chicago: Society of American Archivists, 2017.

²⁶ Federal Agencies Digital Guidelines Initiative. Glossary. <http://www.digitizationguidelines.gov/glossary.php>

and LPs). Digitization standards for video sources are emergent but clear enough that explanatory material can be created that will hold for the next five years or so. An important reason for working with two expert consultants on the project (see below) is to make sure that the most relevant and stable technical standards are highlighted in the modular learning resources.

Project Organization

Principal Investigator: Dr. Paul Conway (Ph.D., Library and Information Studies, University of Michigan, 1991) serves as Principal Investigator for the project. He is associate professor at the University of Michigan School of Information. His research focuses on archival science, the digitization and preservation of cultural heritage resources, and the ethics of new technologies. He teaches two graduate-level courses on the digitization of cultural heritage resources and is an active participant in the review of FADGI standards for still imaging and audio-visual reformatting. Dr. Conway will be responsible for developing and vetting (with experts) all educational and training content for the modules, for revising the products based on advice from the two project consultants, and for arranging for online and print-on-demand publication. He will coordinate technical advice from the two paid consultants and solicit editorial and pedagogical advice from the panel of experts. He will also produce articles for publication that describe the project. [See Proposal Documents for current CV.]

Graduate Student Assistant: Our proposed project will retain a single graduate student from the University of Michigan School of Information on a part time basis (ca. 15 hours per week) for two academic years and one summer.²⁷ This student will be responsible for supporting the research and information assembly needs of the Principal Investigator. The assistant will also maintain a project website and help coordinate the review of modular content by the expert group. The PI has utilized graduate students successfully for support in past research projects. [See position description in Proposal Documents.]

Consultants: Integral to the design of the project is active and ongoing advice from two independent consultants on digitization standards. The PI has worked with both consultants informally on past projects. In particular, both consultants have provided the PI with resources and technical documentation for use in the classroom. Both of the consultants are the very best experts on digitization guidelines promulgated by FADGI in their fields. If the proposed project is approved for funding, the University of Michigan will negotiate formal consulting contracts based on Mellon Foundation guidelines and UM standard contracting procedures. . [See Proposal Documents for current CVs and commitment letters.]

Don Williams, President of Image Science Associates, will provide technical advice on variances in outcomes from scanning equipment calibration as well as the accuracy and precision of measurements of test targets against imaging standards. He will also review the written content of the modules on still image digitization for technical accuracy. Don worked for Kodak for 30 years before retiring and was largely responsible for developing the image testing procedures for digital cameras, most of which now are international standards supported

²⁷ The PI has identified a highly qualified student who will matriculate in the School of Information in fall 2018. If the project is approved, the PI will recruit this student for the duration of the project.

by ISO.

Carl Fleischhauer, an independent consultant recently retired from the Library of Congress, will provide advice on explaining emerging technical standards on the digital conversion of analog video resources. He will also review the modular content for technical accuracy, particularly on issues where technical standards are still in a state of flux. Carl is one of the principle authors of the proposed IASA TC-06 guidelines on video source digitization, so he is particularly well qualified to provide advice on this project. [See Proposal Documents for current CVs and commitment information.]

Expert Advisory/Editorial Board: A cadre of experts, chosen for their deep domain knowledge of one or more components of the project, will guide and inform the work of the project, particularly the written components developed for each module. The primary contribution of the consulting experts is to read and review the written products produced in the proposed project and suggest examples and case studies whose lessons will resonate with a wide audience in the audio-visual components of the modules. All of the following experts have agreed to volunteer their knowledge and experience to the project at no direct cost. The PI will work to maximize the value of this external expertise and minimize the amount of time required to provide advice. It is very important to note that several expert advisers will be advising the project in their roles as part IASA or the FADGI collaborative, rather than in their professional capacities. [See Proposal Documents for current CVs and letters of commitment]

Tanya Clement, Associate Professor, iSchool, University of Texas at Austin, is a digital humanist with a particularly strong focus on extracting content and meaning from audio sources. Tanya has agreed to test the educational resources on audio-visual digitization standards in the classroom at UT Austin and provide feedback on the usefulness of the modules for graduate-level education. Bio: <http://tanyaclement.org/>

Mats Dahlstrom, Associate Professor, Swedish School of Library and Information Science, is an insightful critic of large-scale digitization projects and a proponent of “critical digitization practice” that seeks to surface and examine the agency of digitization processes and practices. Mats has agreed to provide feedback on the structure and intellectual depth of the modular units on still image digitization. He has also agreed to test some of the teaching modules in a graduate level course on still image digitization at the Swedish School of Library and Information Science. Bio: <http://www.hb.se/en/Research/Research-Portal/Researchers/Dahlstrom-Mats/>

Bertram Lyons, Senior Consultant, AVPreserve, is a member of the Executive Committee of IASA. Over the past decade, he has developed tools, policies, and partnerships around the development and management of analog and digital archival collections, particularly in association with AVPreserve. Bert will help ensure that written components of the modules accurately portray the features and functionality of the AVPreserve tools. Bio: <https://www.weareavp.com/people/bertram-lyons/>

Will Prentice, Vice-chair, Training & Education Committee, IASA. In his capacity as a leader of IASA (rather than his professional role as Training & Dissemination Manger within the Sound & Vision Department of the British Library), Will has agreed to connect the proposed project to

IASA's educational initiatives and to sponsor a workshop at the IASA annual conference in Accra, Ghana. The workshop is an opportunity to elicit priorities for training materials on audio-visual resources and to coordinate the proposed project with IASA initiatives. *Wired* Portrait: <http://www.wired.co.uk/article/british-library-sound-archive>

Katherine Wisser, Associate Professor, School of Information and Library Science, Simmons College, is an expert on metadata and its application in digitization and digital curation. Katherine will review the written components of the modules pertaining to technical metadata to help make sure that the modules cast their content at the proper level of complexity for the target audiences. Bio: <http://www.simmons.edu/Faculty/Katherine-Wisser>

Dan Zellner, Production Coordinator, Repository and Digital Curation, Northwestern University Libraries, is a skilled professional who manages digital imaging projects and workflows. He is also the founder and coordinator of the ad hoc FADGI user group. Dan has agreed to test the appropriate digitization workflow modules in the context of staff training and development in digitization labs. He will also connect the project to the work of the DLF Cost Assessment Working Group (particularly the Digitization Cost Calculator). LinkedIn: <https://www.linkedin.com/in/dan-zellner-7b853b/>

Additionally, the Principal Investigator will exercise close coordination with the leadership of FADGI, building on well-established working relationships. In their capacities as employees of the Library of Congress, two leaders in the FADGI initiative cannot endorse the proposal without upper management approval, which would likely delay significantly the start of the project. Quoting from correspondence exchanged between FADGI leadership and the Principal Investigator, "The way you have represented our participation, as representing FADGI, will work. In that capacity we represent a multi-agency Federal initiative, not the Library. As you know, that is an important distinction at this stage. We have a green light to participate." [See Proposal Documents for the full email exchange.]

Products: Modular, Flexible, Self-paced Learning Modules

In developing and publishing learning resources for understanding digitization standards, the project embraces the notion that distinctions are collapsing between flexible and structured online course support content, on the one hand, and the relatively fixed nature of ePub and print-on-demand formats. Both are digital, modular; both now support multimedia content, and both facilitate self-paced, reinforceable, and assessable learning. The proposed project adopts the model of learning support in self-directed online environments, while focusing the content on the application of digitization and technical metadata standards.

In this section, we utilize the following two terms to describe the products that the project will create and deploy.

"Media Domain": A broad area of analog media subject to digitization (along with the focus of attention in the project). In the proposed project, the three media domains (and specific focus) are:

- (1) digital still image reproduction of visual resources (particularly photographic materials);

- (2) digital conversion of analog audio signals (particularly from magnetic tape and grooved media); and
- (3) digital conversion of analog videotape (in the NTSC format most widely used in the United States).

“Learning Module”: A modular, self-contained unit of instruction encompassing a number of components that vary depending on the purpose of the learning module. Each module is built around a single complete idea (e.g., measuring image resolution; size of source material and digital camera capabilities; the structure of a Broadcast Wave File; encoding color in bits and bytes). The module exists as online content that can be easily reproduced as a print-on-demand book or an ePub version. “Learning Modules” appear in the online context as webpages but when printed or downloaded as ePubs take the form of chapters on a book. Each “Learning Module” will contain URLs directing the learner to specific technical standards documents (some of which may exist behind paywalls) and supplemental content. Examples of supplemental content may include an audio stream (podcast), a video stream, a slide presentation, or technical guidelines and documentation.

The proposal describes a two-year (2018-2020), self-contained project (Dig4E – Digitization for Everybody) to create, vet with a group of highly experienced experts and consultants, and then publish highly modularized learning resources that support standards-based digitization of cultural heritage resources (still image, audio, video). We estimate that each media domain will require between 10 and 15 Learning Modules to cover the intellectual territory appropriately. The PI will work with the consultants to determine the precise number of modules for each media domain. See Project Documents for outline of the modules, subject to revision after consultation with the expert group and the consultants.

Each “Learning Module” will feature one or more case studies of good and not-so-good outcomes of digitization process processes, with particular reference to points in a workflow where problems and errors arise due to a combination of material defects, technology system failures, or post digitization processing. As much as possible each “Learning Module” is independent, self-paced, and repeatable. As appropriate two or more modules are clustered into groups by concept, particularly those modules pertaining to the technical standards for digitization promulgated by FADGI and IASA. See Project Documents for an outline of “Learning Modules” for each media domain (still image, analog audio, analog video).

Self-Assessment of Learning: In large scale online learning environments, assessment serves multiple purposes, including establishing pre-requisite knowledge in a “ladder” of learning modules, incentives to proceed or even increase the pace of learning, a demonstration of basic competency, or an assessment of prior learning. For this project, assessment is formative rather than summative in nature. Each “Learning Module” includes a section on self-assessment that provide feedback to the learner on the most important points to remember in any given module. The self-assessment components will be designed for “pencil and paper” completion in a print on demand book or as scored feedback in an online environment. The primary pedagogical goal of the multiple choice quizzes is to give learners a sense of the material covered within each module. Learners would be encouraged to take the quizzes repeatedly to help them understand the material.

Content Production and Distribution

The book is the operative metaphor for the structure and organization of the Learning Modules, each of which will exist simultaneously as freely available web-accessible components and very modestly priced print-on-demand publications in either paper or ePub formats. The web accessible version of the learning modules, available free and directly through the project website: www.Dig4E.org, is the primary focus of content development. The online components include textual explanations of the relevant standards, illustrated as appropriate with graphic material either created by the PI, available through open Creative Commons licensing, or provided by FADGI as public domain content. The online textual and graphic materials (also downloadable in PDF format) will be supplemented by short audio or audiovisual presentations that replicate the essence of the textual material but also enhance the textual presentation with demonstrations of technical processes, such as calibrating a scanner, assessing the physical condition of source materials, operating playback equipment, or working through a workflow document step-by-step. It is our sense that audiovisual supplements to textual materials will enhance accessibility and understandability for learners who may more effectively learn via the spoken word or through demonstrations of techniques.

Based on the experience of colleagues who develop and teach online courses where the content is freely delivered, learners sometimes desire to have a textbook like publication in hand that contains the core content of the learning modules plus URL pointers to the online content. Since the University of Michigan does not sponsor a direct print-on-demand service, the project will work through the CreateSpace service of Amazon to make the textual and visual content of the learning modules available as a completely optional print-on-demand version in trade paperback form and an ePub version suitable for Kindle and other eBook reading devices.²⁸ These versions will be available for purchase through Amazon for \$9.99, with a Kindle version downloadable for \$1.99. CreateSpace estimates that a 100-page black & white, print-on-demand volume priced at \$9.99 will generate a royalty of \$3.84 for each copy sold. Royalties will be retained by the University of Michigan School of Information for use in updating the Learning Modules as standards evolve. In all aspects of the distribution of the learning modules, the project will fully conform to the intellectual property guidelines of the Mellon Foundation.

The structure and pricing of the independently produced publications is modeled on *Introduction to Networking: How the Internet Works*, a course-like introduction to the technical underpinnings of the Internet developed by Charles Severance in 2015.²⁹ In this model, the free and open web version of the book with its flexible navigation through learning modules is accessible through a dedicated website, <http://www.net-intro.com>. The print and ePub versions (\$9.99 and \$1.99, respectively) contain links to the online version, while the online website provides links to download the volume as a PDF document, or purchasing the

²⁸ CreateSpace. <https://www.createspace.com/Products/Book/>

²⁹ Introduction to Networking. https://www.amazon.com/Introduction-Networking-How-Internet-Works/dp/1511654945/ref=sr_1_2?ie=UTF8&qid=1529330234&sr=8-2&keywords=introduction+to+networking+how+the+internet+works

print or e-Pub version through Amazon. The online version allows for the provision of supplemental materials in audiovisual formats, the automatic scoring of quizzes that accompany each Learning Module, and hyperlinks to glossary terms.

All of the content for the learning modules will be created directly by the Principal Investigator with production support from a graduate student assistant. All of the text will be composed in the Markdown text processing platform, which allows for the simplest and most portable composition of text. Using the universal document converter Pandoc,³⁰ the basic text can then be transformed into other text-based files, such as HTML, LaTeX, or MS Word as needed. Using Markdown in combination with Pandoc will facilitate the international review of the Learning Modules by the expert advisors as well as the efficient publication of the modules in web-accessible and print-on-demand formats.

This audio-visual components of the project will employ a simple configuration of medium grade recording tools (Rode Procaster microphone; Panasonic HC-VX1 camera) and commercial capture and editing software (Camtasia) in a modified DIY model based in an informal recording studio environment in the School of Information. The DIY approach has and continues to yield excellent quality results with low overhead costs.

In terms of training and production support, the project team will be able to take full advantage of the University of Michigan's campus-wide efforts to adapt large-scale online learning methodologies to transform the residential learning experience and to launch new online degree programs. Through the Office of Academic Innovation (OAI), Michigan has established a robust infrastructure to train and support faculty who are creating digital learning content and delivering that content through local infrastructures and a mix of commercial and open access partnerships.³¹ Most recently in partnership with Coursera and OAI, the School of Information has announced plans to launch a new online-only degree program in applied data science.³² The proposed project leverages the School's newly defined faculty support model for developing, and sustaining online teaching resources. The PI has enthusiastic backing from the School of Information for the proposed project because of its relevance to the School's strategic priorities for the online distribution of learning resources.

Schedule of Major Activities

The two-year project will run from October 1, 2018 through September 30, 2020. The project is largely administered by the Principal Investigator (PI) with the part-time support of a graduate student assistant, the committed advice of an Expert Advisory Group, and two paid consultants. The rhythm of the project involves drafting text and multimedia content into modules, with the help of two world-class consultants, thoroughly vetting the content with a group of domain experts, coordinating the structure of the modules with the leadership of FADGI on a regular basis, revising as appropriate, and then publishing the modules online and in a print-on-demand formats. The website www.dig4e.org will serve as an open and publically accessible portal for

³⁰ Pandoc: A Universal Document Converter. <http://pandoc.org/>

³¹ University of Michigan. Office of Academic Innovation. <http://ai.umich.edu/about-ai/>

³² Coursera. Master of Applied Data Science. <https://www.coursera.org/degrees/master-of-applied-data-science-umich>

the online learning modules, the content of which will also be preserved in the University of Michigan's DeepBlue Data repository.

Fall 2018 (October to December)

- The PI will start up the project, including recruiting of graduate student assistant and creation of project website (Dig4E.org). The student assistant will do most of the work to create the website, which will be partially built on the University's supported WordPress platform, with links to modular learning resources stored on a University supported server.
- The PI and graduate student assistant will begin assembling first-round documentation for digital still image and audio conversion modules, including standards documents, and secondary literature.
- The PI will finalize a detailed list of modules for each media domain (still image, audio, and video), the relationships among the individual modules, and the types of component content appropriate for each module. . [See Project Documents for an outline of the learning modules for each of the three media domains.]
- The PI, with the help of the two consultants, will finalize the detailed learning objectives for each of the individual learning modules. Based on these objectives, the PI and graduate student assistant will create self-assessment quizzes (typically multiple choice questions) for the digital still image modules.
- The PI will develop the written and visual supplements for all learning modules for digital still image standards.
- Consultant Don Williams will review written content and provide visual examples of "less than optimal" outcomes test target evaluations
- The PI travels to Washington DC for three days to meet with the leadership of FADGI, based administratively at the Library of Congress, and representatives of up to 19 other government agencies that are part of the FADGI collaborative, including the Smithsonian Institution, the National Archives and Records Administration, and the Government Printing Office. Because so much FADGI activity happens in Washington, DC, it is most effective for the PI to travel there with scheduled meetings. This work of active coordination with FADGI is a critically important part of the project. The PI will travel to Washington four times over the course of the two-year project, once each academic term.

Winter 2019 (January to May)

- The PI, with the help of the two consultants, will finalize the detailed learning objectives for individual learning modules for audio and video standards. Based on these objectives, the PI and graduate student assistant will create self-assessment quizzes (typically multiple choice questions) for the digital still image modules.
- The PI and the two consultants will work together to identify the specific case study documentary evidence for all learning modules and to identify likely candidates.
- Selected members of the expert advisor group will conduct an editorial review of digital still image modules and the audio standards learning modules.
- The PI will finalize and record multimedia components of still image standards modules.
- The PI will finalize the audio-visual supplemental content for the analog videotape conversion learning modules.
- Travel to Washington DC to work with representatives of the FADGI community. See fall

2018 schedule for rationale.

Summer 2019 (June to August)

- The PI will arrange for publication of the digital still image learning modules, likely via Amazon Publishing Services for print-on-demand and ePub formats. The website Dig4E.org will serve as a portal for the online versions of the modules, including self-study quizzes.
- The PI will continue with development of learning modules for audio standards modules, including text, audio-video supplements, and case study evidence. Consultant Carl Fleischhauer will review the structure and flow of the audio and video modules.
- The PI, with the assistance of the graduate student assistant, will create and publish for the general public all audio digitization standards training modules, likely via Amazon Publishing Services for print-on-demand and ePub formats. The website Dig4E.org will serve as a portal for the online versions of the modules, including self-study quizzes.
- The PI and the graduate student assistant will develop first draft content for analog videotape modules, including text, audio-video supplements, and case study evidence. Consultant Carl Fleischhauer will provide technical advice on the structure and content of the module content.
- Selected members of the expert group will provide a thorough editorial review of the content of the analog videotape modules.

Fall 2019 (September to December)

- The PI will pilot the still image digitization modules in UMSI graduate course; revise assessments after term ends. Expert Adviser Mats Dahlstrom will use the learning modules for still image digitization in his graduate course at the Swedish School of Library and Information Science.
- Selected members of the expert adviser group will conduct an editorial review of learning module content for analog videotape conversion standards.
- The PI will finalize the content of the analog video modules, including text, audio-video supplements, and case study evidence.
- The PI will pilot audio and videotape modules in a UMSI graduate course and make revisions to the content based on student assessments and feedback.
- The PI will travel to Washington DC to work with representatives of the FADGI community. See fall 2018 schedule for rationale.

Winter 2020 (January to May)

- The PI will publish to the general public analog video modules, likely via Amazon Publishing Services for print-on-demand and ePub formats. The website Dig4E.org will serve as a portal for the online versions of the modules, including self-study quizzes.
- The PI will work with the members of the expert advisory group to complete a full review of all module content. Such coordination will take place via videoconferences with live access to all digital content. Some of these review conferences will be one-on-one with individual experts, while others will feature a round robin discussion with the entire advisory group.
- The PI will travel to Washington DC to work with representatives of the FADGI community. See fall 2018 schedule for rationale.

- The PI will make revisions to the content of all modules based on review and feedback from the expert advisory group.

Summer 2020 (June and August)

- The PI will complete documentation on all learning modules
- The PI will coordinate the marketing of all learning modules. Both FADGI and IASA have agreed to provide a link to the learning modules on their official websites. The Society of American Archivists has agreed to list the learning resources in their directory of archival education – initially associated with the offerings of the University of Michigan. In publishing the print-on-demand and ePub versions of the learning modules through Amazon Publishing Services, there is a significant amount of built in publicity through Amazon. Finally, the Northeast Document Conservation Center has agreed to promote the learning modules through their website (which contains a significant number of relevant but somewhat outdated pamphlets on digitization) and as part of their nationwide series of educational conferences. It is likely that the PI will attend and promote the learning modules at the annual Society of American Archivists meeting (Austin, TX). Funds to support this activity are not requested from the Foundation.

Fall 2020 (September)

- The PI will complete a full report on the project.
- The PI will produce an article for a peer reviewed journal (*Library Quarterly*) that explains the underlying concepts of the learning modules, including self-assessment.

Outcomes and Benefits

The project is ambitious in terms of the quantity, variety, and diversity of content that will be created and published to support learning about digitization standards. In contrast, the project is modest in its claims for immediate and wide impact, particularly in the course of the two-year project. We aspire to offer some really well developed and well-vetted learning materials to help practitioners across a range of audiences (identified above) understand the value and present maturity of digitization standards and apply them to create well-formed digital surrogates of some of the most common cultural heritage resources. The best impact that we can expect in the short run may be in measuring and counting and modestly assessing the use of the objects. But in the two year life of the project, we mainly hope to develop and vet the learning materials, publish them for free and open use, promote their use, and establish the mechanism for tracking use.

The primary products created in Dig4E – Digitization for Everybody are modular learning resources (online and in print-on-demand formats) that support standards-based digitization of cultural heritage resources (still image, audio, and video). We envision that the modules will serve collectively as a freely available toolkit for mastering existing and emerging technical standards for digitization processes, workflows, and technical metadata, along with critically important software tools for quality assessment and digitization priority setting. The proposed learning objects will have a very high level of credibility because the project will tap the expertise of two independent consultants on digitization standards and a group of scholars and expert professionals to help insure technical accuracy and readability. The involvement of

experts and the iterative development of the content of the modules will help ensure that the intended audiences will grasp and assess their understanding of the content.

Sustainability and Reuse

All content will be freely available and distributed through the project website Dig4E.org under a Creative Commons BY-NC-SA license, which lets others use, revise, and build upon the work non-commercially, as long as they attribute their work properly and release any revisions under an identical CC license. This type of license maximizes reuse capabilities while preserving the openness of the content.

The University of Michigan will preserve all educational content and any associated computer code in its open digital repository, *Deep Blue Data*, which is the university's service for preserving digital research data built upon the Samvera/Fedora platform. As a tenured member of the University of Michigan faculty, the PI has permanent and unlimited use of *Deep Blue Data*.

The Learning Modules will function for stand-alone use, through the website (with links to print-on-demand and ePub formats, in formal courses, professional education workshops, or on-the-job training programs. Royalties earned on the sale of print-on-demand materials will be retained by the University of Michigan for use in updating the Learning Modules as standards evolve. We are committed to maintaining the accuracy of the materials for at least five years from the point of publication.