Avalon Media System: Integrating audiovisual collections with research, digital preservation, and a sustainable developer community

Abstract

Northwestern University Libraries, in partnership with subcontractor Indiana University Libraries, and working with contractors/consultants Lyrasis and Artefactual Systems, propose a National Digital Platform project to bolster sustainability of the Avalon Media System open source audio and video access repository system while increasing its value to current and future library adopters. This two- year project would run from July 1, 2017 – June 30, 2019.

Avalon is an open source software package for managing and providing access to large collections of audio and video materials, currently fully implemented at five institutions. Planned, designed, and built with support from prior grants from IMLS and the Andrew W. Mellon Foundation, Avalon provides to library, archives, and cultural heritage institutions what existing commercial, open source media systems, and digital repository platforms do not adequately address: an affordable, community driven, preservation ready audio and video (AV) delivery system. However, we have seen that many institutions have needs for media access but lack the resources to locally stand up an instance of a media-focused repository system such as Avalon. In addition, institutions want to be able to integrate more sophisticated AV functionality into existing repository systems such as institutional and research data repositories alongside other content formats.

The aim of this project is to increase adoption of Avalon within the library and archives community by enhancing the value proposition of Avalon and carrying out work to help ensure sustainability:

1. **Integrate Avalon within the Hydra community.** We will adapt Avalon to make use of the current Hydra open source repository software stack and will engage members of the Hydra community in development through community sprints. We will also increase the modularity of Avalon so that its components can be more easily integrated into other Hydra-based repository systems, including Hydra-in-a-Box, to provide time-based media access.
2. **Implement a cloud-hosted Software-as-a-Service (SaaS) version of Avalon.** We will work with Lyrasis and DuraSpace on the establishment of a hosted SaaS offering by conducting pilots to help provide cost modeling for Avalon as a cloud based service. If successful, revenues from the service are anticipated to help to sustain future Avalon development and maintenance.
3. **Connect Avalon to media preservation systems and workflows.** We will connect Avalon to long-term digital preservation solutions and improve interoperability with workflow and management tools in order to provide a complete AV preservation and access platform, by working with Artefactual Systems to integrate Avalon with the Archivematica platform.
4. **Enable interoperability with scholarly tools.** The Avalon technical team will take a leadership role with the International Image Interoperability Framework (IIIF) by contributing to the creation of an AV interoperability spec and providing a demonstration implementation.

The functional enhancements to Avalon in the areas of preservation and interoperability with scholarly tools and the support for sustainability generated through SaaS implementation and greater development engagement within the Hydra community will improve the ability of libraries and archives to integrate robust AV media delivery into their digital collections practices and increase the AV collections available as part of the national digital library platform.

Avalon Media System: Integrating audiovisual collections with research, digital preservation, and a sustainable developer community

1. Statement of National Need

Audio and video collections have become an increasing area of focus for academic and research libraries and their users in recent years. This is due to a number of factors, including an increased focus on special collections; increased demand for media materials from students, instructors, and researchers; a recognition of the limited time window for reformatting of analog audio and video formats due to issues of degradation and obsolescence; and decreases in storage and network bandwidth costs, which make digitization, storage, and delivery of large audio and video files more feasible and cost-effective. This increased national focus on audio and video (AV) is also evidenced through funders' attention to this area, including grants made by NEH, the Andrew W. Mellon Foundation, and others to support reformatting of and access to specific collections, and the Recordings at Risk program recently announced by the Council on Library and Information Resources, supported by the Mellon Foundation. Despite this increased focus, AV materials are not well represented in the emerging national digital platform, in part due to the difficulties that remain for libraries to successfully execute digitization and access programs for these formats. As an example, as of January 1, 2017, audio and moving image items represent only 0.5% of the content of the Digital Public Library of America (DPLA).[1]

One specific area of need that Northwestern University (NU) and Indiana University (IU) have worked to address over the past six years is that of platforms for providing online access to AV collections, through the development of the Avalon Media System. Avalon is an open source software package for managing and providing access to large collections of audio and video materials. Planned, designed, and built with support from prior grants from IMLS (planning 2011-2012, implementation 2012-2016) and the Andrew W. Mellon Foundation (2015-present), Avalon provides to library, archives, and cultural heritage institutions what existing commercial, open source media systems, and digital repository platforms do not adequately address: an affordable, community driven, preservation ready AV delivery system.

Many commercial products and services—including media streaming and delivery servers, digital asset management systems, classroom capture systems, and cloud-based online video environments—exist to support various forms of digital audio and video distribution, but they generally have not been flexible enough to support the full set of online media management and delivery needs of libraries and archives because they are tailored to other use cases. As primarily closed source systems, they can also not easily be adapted to integrate with existing library repository systems and workflows. Support for deep description in the form of structural or time-based metadata, robust authentication and authorization, and integration with repositories providing services for long-term preservation are key areas where these tools most often fall short. Avalon is the only open tool that truly serves the needs of academic libraries, archives, and the higher education enterprise for managing access to streamed video and audio collections and enabling their use by students and faculty.

---

[1] Based on facet counts at http://dp.la/, 1 January 2017

Avalon has been developed to support a wide range of use cases for online audio and video access, including special collections, licensed AV collections, and instructional collections. As a result, it features flexible access controls, allowing access to be limited at either an item or collection level to individual users, groups, IP address ranges, or course rosters based on copyright, licensing, and fair use concerns. It also provides collection managers and other users the ability to easily embed hosted audio and video into external websites, web content management systems, learning management systems, and online exhibits.

Avalon has been fully implemented at five institutions: Northwestern University, Indiana University, University of Virginia, Washington University, University of Alberta, and Calvin College, and a number of other institutions are in various stages of evaluating, testing, and putting Avalon into production, including Yale University. At least one vendor is currently offering installation and support of Avalon.[2] Based on our market research and discussions with prospective adopters, we have seen that many institutions have needs for media access but lack the resources to locally stand up an instance of a media-focused repository system such as Avalon. A recent survey conducted by the Avalon team and Lyrasis showed at least 31 institutions potentially interested in a cloud-hosted Software-as-a-Service offering of Avalon.[3] In addition, some institutions want to be able to integrate more sophisticated AV functionality into existing repository systems such as institutional and research data repositories alongside other content formats.

These findings have been validated by the IMLS-supported Hydra-in-a-Box project's recent findings that time-based (AV) media access is a major need of libraries for repository systems.[4] However, audio and video are not a focus of the Hydra-in-a-Box project's current IMLS-supported development project.

In this project, NU and IU propose to build on Avalon's track record of success to improve interoperability with other Hydra-based repository applications and digital preservation tools, enable interoperability with emerging standards-based tools for scholarly use of online audio and video, and establish broader library community engagement in both the use and development of Avalon by implementing a cloud-hosted Software-as-a-Service version of Avalon and conducting community development sprints.

2. Project Design

The overarching aim of the proposed project is to increase adoption of Avalon Media System and increase online availability of audio and video collections from libraries and archives by 1) enhancing the functionality and value proposition of Avalon and 2) carrying out work to help ensure sustainability of the Avalon system into the future.

Through user feedback, discussions with the Hydra-in-a-Box project, the success of projects like the International Image Interoperability Framework (IIIF) for images and newspapers, and the increased

---

[2] See https://curationexperts.com/avalon/

[3] See survey report in supporting documents

[4] See https://wiki.duraspace.org/display/hydra/Hydra-in-a-Box+Design+Documents?preview=/76841397/76843894/Hydra-in-a-Box%20User%20Survey.pdf, p. 24

attention being paid to preservation of digital materials, as well as a recent survey of potential Avalon users, four major goals presented themselves for the focus of this grant.

Project Goals

Goal 1: Implement a cloud-hosted version of Avalon

Building off technical work and partnership discussions begun under our current Mellon grant, IU and NU will work with Lyrasis and DuraSpace to pilot test a cloud-hosted Software-as-a-Service (SaaS) offering of Avalon and to develop cost and partnership models for an ongoing service. Running in Amazon Web Services, and integrated with DuraSpace's DuraCloud service for long-term storage of master files, the Avalon SaaS offering is expected to be an appealing option for audio and video access for smaller institutions with minimal technical staff or for medium-to-large institutions for whom AV repository infrastructure is not a core focus.

IU and NU will work with Lyrasis to refine Avalon's deployability and manageability in the AWS cloud environment and to identify approximately ten institutions to serve as participants in a six-month Avalon SaaS pilot test to be managed by Lyrasis, with infrastructure costs and staffing supported in part from IMLS funding for the proposed project. These tests will provide IU, NU, and Lyrasis with real numbers on the cost of installation, maintenance, and technical support, and any additional development needed to support scaling Avalon as a SaaS product. Any enhancements made to Avalon would become part of the open source product, open to other users and service providers via GitHub. Assuming the pilot is successful, IU and NU will help Lyrasis establish an ongoing Avalon SaaS offering.

Goal 2: Integrate Avalon within the Hydra community

NU and IU will work to adapt Avalon to make use of the current version of the Hydra stack and to modularize Avalon's code by engaging members of the Hydra community beyond NU and IU to identify areas of functionality useful for other Hydra-based repository applications. Avalon development team members will work with other Hydra developers to break apart these functions into modules, or "gems" in Ruby programming language terminology, that can more easily be reused across applications and can be maintained and enhanced by the broader Hydra developer community, not just developers at IU and NU.

This work will also allow Avalon's AV-related capabilities to be presented in other Hydra-based applications, so that institutions with needs for managing and delivering AV materials alongside other formats in the context of research data or digital collections repositories may do so. The Avalon team will work closely with other major Hydra-based projects, including Hydra-in-a-Box, to make sure that features are broken out in a way that is useful to them.

A major method of engaging developers beyond the existing Avalon team will be through the use of *community sprints* that solicit participation from developers and institutions across the wider Hydra community, a process that has proven successful in other Hydra projects such as Sufia.

Goal 3: Connect Avalon to media preservation systems and workflows

As Avalon is focused on access concerns, many institutions will want to connect it into workflows and systems being used for long-term digital preservation activities. Current Avalon implementations at IU and elsewhere have achieved this through custom coding, but more generally reusable preservation mechanisms would enhance the attractiveness and usefulness of Avalon to a broader set of potential adopters. Work on the SaaS offering discussed above will help in part by allowing Avalon to support access to content for which master files are stored in DuraCloud, which could be migrated to DPN (Digital Preservation Network) via the DuraCloud Vault and Chronopolis services.

However, beyond long-term file storage, institutions also need support for activities such as file characterization and preservation metadata extraction for digitized and born-digital AV materials. The project team will work with Artefactual Systems on integrating pieces of their open source tool, Archivematica, into Avalon ingestion workflows to provide preservation metadata and better packaging for long-term digital preservation activities.

Goal 4: Enable interoperability with scholarly tools

In addition to providing effective tools for managing access to digital AV collections, Avalon has always had a goal of providing usable and useful functionality for collection users, particularly students, teachers, and scholars. Avalon's predecessor, Variations,[5] provided a deep set of music-focused pedagogical and scholarly functionality, including tools for annotation, diagramming, audio-score synchronization, and testing/quizzing, that were tightly integrated with the underlying delivery platform. While this made users very happy, this model of tightly-coupled UI tools and repository platform did not allow for easy maintainability of tools or easy application of the underlying delivery system to other types of AV content.

Over the past several years, the International Image Interoperability Framework (IIIF) has emerged as a successful standards effort in integrating scholarly and teaching and learning user interface tools with a wide variety of still image content repositories. Supported by both open source and commercial platforms, IIIF is a clear demonstration of the value of standards in enabling innovation.

A new effort, known as IIIF A/V, is now underway to extend the IIIF standards to encompass audio and video content and repositories, and the Avalon team has been actively engaged in the process. This project would support work by the Avalon development team to remain actively engaged in the development of the IIIF A/V standards and to implement support for them within Avalon, allowing AV content delivered by Avalon to be used by emerging IIIF-compliant tools to create annotations and presentations for scholarly and teaching purposes.

In conjunction with the reuse enabled by IIIF A/V, this project will also include implementation of a Rights field for digital objects in Avalon with preconfigured options for Creative Commons licenses or RightsStatements.org statements. This field will pre-populate with linked open data URIs that will be added to the metadata for the digital objects. The entry form will allow for Rights fields to be assigned at varied levels of the time-based media with "more information" tooltip bubbles provided in the interface for basic guidance, as time-based media often has complicated rights issues at multiple levels. Work on implementation of a Rights field will be carried out in consultation with the

[5] http://variations.indiana.edu/

Penn State University Libraries, which will provide in-kind support for creation of guidance materials by the Penn State Libraries Copyright Officer, Brandy Karl, who is an experienced Fair Use litigator, in conjunction with a Penn State instructional designer. The RightsStatements.org URIs will enable users to identify rights for use of the objects, as well as allowing machine-readable statements consistent with data from DPLA and Europeana. Providing this feature in Avalon enables Avalon users to comply with the Rights recommendations of these aggregators.

Potential Risks

In terms of risks to the project, there are a few. In its current iteration, Avalon does not have the technical documentation necessary to immediately start community sprints. Work has already begun to enhance documentation, and a sprint with developers outside the current core group of Avalon developers, is being scheduled for early Spring 2017 to help identify any gaps in information and resources that will need to be addressed before successfully running community sprints for Avalon integration with Hydra. Additionally, loss of staff could cause timeline slips, however both institutions have successfully dealt with significant staff turnover during previous phases of Avalon development and feel confident in the contingency plans in place and shared expertise across project staff to mitigate risk.

The biggest risk, however, is not finding the best candidates for the Community Technical Lead and the Community and Project Manager positions. Without the right candidates to implement the vision of the project directors, it may take longer than we hope to gather community involvement in the Hydra integration sprints, the future planning of the project, and the development of things like the IIIF framework. Ideally, the successful candidates will come from the Hydra community already and won't need as much onboarding in terms of learning about the community. However, if we do end up hiring a candidate from outside the community, both institutions have solid ties to the Hydra community and feel that if they are unable to hire someone already familiar with Hydra, they will be able to bring them on board relatively quickly. It may take longer than we'd prefer to do this, though, which could cause some slippage in deliverable timelines. As always, salaries are a concern, but we feel confident that the salaries requested should provide us with a pool of good candidates.

Project Activities

Work on the project will be broken down into four major overlapping phases, which are further detailed in the attached Schedule of Completion:

Phase 1: Startup (July-September 2017)

During this time, the project directors will post and hire for the two new positions to be funded by the grant: Community and Project Manager and Community Technical Lead. The project directors and the Community and Project Manager, once hired, will also work with Lyrasis to identify and select institutions to participate in the SaaS pilot, and the DevOps Engineer will work with Lyrasis and DuraSpace to complete any technical work necessary to be ready for the SaaS pilot.

Phase 2: Pilots and Design (September 2017 – May 2018)

In phase 2, the SaaS pilots will be conducted, and the Community and Project Manager and Community Technical Lead will work with members of the Hydra community to identify functionality for modularization and shared development sprinting. One or two community sprints will be carried out. The Community Technical Lead and other project team members will participate in IIIF AV spec

development. Initial conversations about creation of rightsstatments.org guidance materials and their implementation will begin. Requirements and specifications for Avalon-Archivematica interoperability will also be developed.

Phase 3: Development and Implementation (April 2018 – March 2019)
Development of Archivematica functionality will be carried out by Artefactual, and the Avalon development team will carry out any necessary changes to Avalon to support this integration. SaaS pilots will be completed, and cost analysis and recommendations for technical enhancements will be developed and used to transition the pilot into an operational service. Additional community sprints on modularization with other Hydra community developers will be conducted. Preliminary support for the IIIF AV spec in Avalon and a proof-of-concept IIIF AV client will be developed. Rightsstatments.org guidance materials will have been tested.

Phase 4: Wrap-up (April-June 2019)
Development on code modularization, IIIF AV, and SaaS pilot will be completed, final documentation will be written, Rightsstatments.org materials will be completed, and structure for post-grant Avalon project management and development will be defined and put in place.

During all four phases, the existing Avalon development team will continue to work on enhancements and bug fixes as dictated by the needs of existing Avalon users and what we know about needs of prospective adopters (SaaS or otherwise) via discussions and surveys.[6]

Project Audience and Input
The primary audience for this project is academic and research libraries and archives with needs for delivery and preservation of digital audio and video collections, along with their users. The project's work plan and staffing structure directly facilitate input from a broad set of stakeholders and potential adopters around both functional and technical concerns, particularly through the Community and Project Manager and Community Technical Lead positions. These positions are designed to build community and gather information and data from the larger community to help guide the progress and the planning of the project. By offering Avalon as a cloud based service, institutions who serve underserved groups and communities, and don't necessarily have the technical staff to install a local implementation or contribute developers to open source projects, will be able to serve their constituencies' audio/video streaming needs while also contributing to the future success of the project.

Resources and Management Plan
The project, which will run for two years from July 1, 2017 through June 30, 2019, will be managed by the Northwestern University Libraries, under the direction of **Evviva Weinraub**, Associate University Librarian for Digital Strategies, as Project Director. The project will be carried out in close collaboration with staff from the Indiana University Bloomington Libraries as a subcontractor, with **Jon Dunn**, Assistant Dean for Library Technologies, serving as Project Director for IU.

Northwestern University Libraries has a history of pioneering use of technology to support teaching and learning. It was one of the first American research libraries to develop an electronic reserves

[6] See Avalon release road map, https://wiki.dlib.indiana.edu/display/VarVideo/Avalon+Release+Road+Map

service, offering digitized course materials starting in 1994. Northwestern's first online video collection, The Video Encyclopedia of the Twentieth Century, was mounted in 1999 and was retired recently, with plans to migrate the content to its local instance of Avalon. In 2001, the University Libraries actively partnered with School of Music faculty to begin streaming music collections for course use, and added a video electronic reserve service in 2003.

Northwestern is also an active participant in open source software development. The Book Workflow Interface software, a Fedora-integrated workflow management and publishing solution for digitized books, was developed in 2009 under the sponsorship of the Andrew W. Mellon Foundation. In addition to work on the Avalon Media System, Northwestern developed and released the Digital Image Library Hydra stack application, now in production at Northwestern and used to provide access to 115,000 teaching images, including several unique collections created by Northwestern faculty and from the Libraries' special and archival collections.

The Indiana University Libraries have an extensive history in the development of open source software and the use of digital and Internet technologies for preservation of and access to audio and video media. IU's Variations project, launched in 1996, was one of the world's first streaming media digital library systems. Developed and enhanced under support from NSF, NEH, and IMLS, and released as open source in 2009, Variations eventually came to be used at over a dozen colleges and universities beyond IU.

IU's Media Digitization and Preservation Initiative[7] is a comprehensive effort to digitally preserve 280,000+ rare and unique items from audio and video collections across the university, supported by $15 million in internal funding from IU's administration. IU is using Avalon as a key part of its infrastructure to support both internal and public access to collections digitized as part of MDPI.

NU and IU are both partners in the Hydra Project, a growing collaboration of institutions working to develop repository applications based on a Ruby on Rails framework, with Jon Dunn of IU serving as a member of the Hydra Steering Group. NU and IU are both sponsors of the Fedora repository project through the DuraSpace organization, with Evviva Weinraub of NU serving on the Fedora Steering Group and both Evviva Weinraub and Jon Dunn serving on the Fedora Leadership Group.

NU and IU have been successfully co-developing Avalon since 2011, with support from IMLS and Mellon, as well as substantial internal resources from the two partners. Many existing Avalon team members funded by NU and IU will continue in their roles, supplemented by two new temporary positions and one current temporary position to be continued with funding requested from IMLS:

The **Community and Project Manager**, based at NU, will be responsible for coordinating work between the Avalon team and the broader Hydra community structure and Avalon user community. The person in this role will be responsible for gathering user needs from the community and coordinating the SaaS pilot with Lyrasis and DuraSpace and will work on the creation of a sustainability plan for Avalon as a SaaS, training material, and documentation to make implementation and community participation easier. This position will also serve as overall Project

[7] http://mdpi.iu.edu/

Manager for the grant project, tracking project progress and ensuring objectives, promised deliverables, and reporting deadlines are met.

The **Community Technical Lead**, based at IU, will build and expand on the work of existing team members to coordinate community technical efforts around the design and development of Avalon. The person in this role will serve as the project's lead on IIIF AV spec creation and implementation and will coordinate the integration of Avalon with the current Hydra stack (including conducting shared sprints) and integration of AV functionality into HyBox and other Hydra-based applications. They will also provide technical coordination and leadership for work on integration with Archivematica.

Phuong Dinh, DevOps Engineer at IU, will work with Lyrasis and DuraSpace on technical support for the SaaS pilot, packaging, and other related activities.

In addition, funding is requested to support a contract with Lyrasis to fund Amazon Web Services and DuraCloud infrastructure, as well as technical personnel necessary for carrying out the Avalon SaaS pilot. Funding is also requested to engage Artefactual Systems as a contractor to develop integration with Archivematica, an open source digital preservation platform, and with the Archivematica Format Policy Registry, which supplies format specific rules for format identification, validation, and normalization to Archivematica and other open source digital preservation tools.

Travel funding is requested for both NU and IU to support promotion of Avalon and dissemination of project work at key conferences and events, to support participation in relevant Hydra community technical meetings, and to support biannual face-to-face project team meetings, alternating between Evanston, IL and Bloomington, IN. Travel funding is also requested for Penn State Libraries Copyright Officer to participate in two of these project team meetings to assist with implementation.

Core Avalon software development work will continue to be carried out through the Agile Scrum process that NU and IU have successfully used over the past four years, with a single Scrum team distributed across the two sites. Team members communicate regularly using audio and video conferencing, Slack, and email, with brief "standup" meetings every morning and bi-weekly sprint review and planning sessions.

Maria Whitaker (IU) will serve as Scrum Master, responsible for overseeing development team process and trying to remove roadblocks that occur, and **Jon Cameron** (IU) and **David Schober** (NU) will serve as Co-Product Owners, responsible for finalizing and prioritizing user stories to drive development. Existing software developers at both NU and IU will be cost-shared to perform technical design and development work, commitments for which are further detailed in the Budget Justification. NU and IU will also contribute additional staff time as needed to support user experience design and metadata consulting for the project.

Relevant team members will engage in regular calls and email discussions with partner staff at Lyrasis and DuraSpace, and with staff of Artefactual during the phases of the project involving integration work with Archivematica.

Evaluation and Performance Measurement

Tracking of project progress toward achievement of the intended results will be overseen by the Community and Project Manager, and a number of data points will be used to evaluate project success, including: 1) completion of project deliverables and goals as listed in this proposal; 2) number of new Avalon adopters over the course of the project; 3) satisfaction of Avalon SaaS pilot participants as measured by feedback surveys; 4) number of SaaS pilot participants continuing from pilot to paid service; 5) number of outside institutions and personnel contributing code to Avalon development through shared sprints; and 6) number of presentations on Avalon-related activities at Hydra Connect, Open Repositories, CNI, AMIA, and other relevant conferences.

Sustainability

One of the major goals of this grant is to help place Avalon on a sustainable footing by shifting it from a stand-alone project to one more integrated with the larger Hydra community, with development and maintenance responsibility leveraged more broadly across the community via regular community sprints. In addition, NU and IU will work with Lyrasis to explore revenue or developer resource sharing models that allow the hosted SaaS offering to help support ongoing product development and maintenance. Through these efforts and continued contributions from NU and IU, we expect core Avalon code and functionality to be sustainable past the end of the project without ongoing continued grant funding.

3. National Impact

The primary national impact of this project will be to make it easier for libraries to provide online access to their audio and video collections at a time when both libraries and their users have an increased interest in these formats. This work builds on a strong history of leadership and commitment at both NU and IU in the development of open source repository and audiovisual collections tools.

By the end of the project, libraries will be able to take advantage of the open source Avalon platform through a number of different avenues: 1) The integration of Avalon within the larger Hydra stack will allow institutions who are running their own instances of Hydra-based repository systems to deliver AV materials within a single repository platform instead of having to run it as a stand-alone system. 2) This work will also allow projects such as Hydra-in-a-Box to more easily add AV support to their own applications, providing another path for use of Avalon's AV features. 3) A hosted SaaS offering of Avalon from Lyrasis combined with DuraCloud will allow institutions with significant AV access and preservation needs to work with a trusted nonprofit vendor to provide enhanced services. 4) Finally, institutions that choose to continue to run Avalon locally will benefit from increased functionality developed over the course of the project and the fact that Avalon has been placed on a more solid footing for sustainability. In particular, the integration work with Hydra, along with the Community and Project Manager position, will help ensure the Avalon user community's voice is represented as part of the larger Hydra community and that much of Avalon's code is considered to be part of the overall Hydra stack thereby helping to increase sustainability for the project

Additionally, Avalon will be made more useful to both libraries and their users through new technical integrations, and thus more attractive to potential adopters. Integration of preservation workflows using the open source Archivematica tool will facilitate the capture of preservation metadata and deposit into local or cloud-based digital preservation environments such as HydraDAM2, DPN, DuraCloud, or the Academic Preservation Trust (APTrust). Addition of support for IIIF A/V standards

will allow content hosted in Avalon to be annotated, combined and arranged into new forms for scholarship, teaching, and learning, through a variety of external tools.

In reference to IMLS Performance Goals, this project will have a national impact in the following areas:

- *Broaden access and expand use of the Nation's content and collections*, by making it easier for libraries to provide online access to digitize and born-digital audio and video collections through multiple pathways for use of Avalon. Success will be measured by having at least three institutions subscribe to a production Avalon SaaS offering by the end of the grant period. Through surveys of the Hydra and Avalon user communities, we will ascertain the number of planned adopters of Avalon and products using code that originated in Avalon within two years of the completion of the grant period. Avalon on the technical roadmaps of five institutions will be considered success, and ten institutions will be considered a reach goal.

- *Improve preservation, conservation, and care of the Nation's content and collections*, by connecting a robust audio/video access platform to national digital preservation infrastructure components such as DuraCloud and DPN. Success will be measured by surveying the Hydra and Avalon user communities to count the number of institutions planning to use Avalon as a component of an AV preservation effort, with a goal of at least five institutions, as well as a trial ingestion of an AV collection into a dark archival storage solution having been conducted by NU and IU.

Performance results will be reported in interim and final reports to IMLS and via the Avalon project website.

The main products of the project will be: 1) an enhanced version of Avalon suitable for cloud or local deployment, with support for IIIF AV and integration with Archivematica; 2) a sustainable Avalon open source codebase that is well integrated with the Hydra stack and can be easily integrated into other Hydra-based applications; 3) a cloud hosted Software-as-a-Service Avalon offering and accompanying cost model and 4) strong developer and user documentation;

At its core, the primary goal of this grant is to make Avalon a sustainable project. Currently it is primarily developed at Indiana University and at Northwestern University and has been well supported and funded through IMLS and the Mellon Foundation. This grant is designed to make Avalon sustainable by tying it closely with an existing, well supported, open source community active in the galleries, libraries, archives, and museum space to provide the maximum number of outlets for possible adoption. In addition, by working with existing and respected vendors and nonprofits in the library space to provide a cloud based version of Avalon, we are ensuring that a larger number of institutions can adopt Avalon for their streaming and preservation needs which providing a home for the standalone version of Avalon.

Activities Year 1	July 2017	Aug. 2017	Sept. 2017	Oct. 2017	Nov. 2017	Dec. 2017	Jan. 2018	Feb. 2018	March 2018	April 2018	May 2018	June 2018
Hire two new positions to be funded by the grant: Community and Project Manager; Community Technical Lead												
Select institutions and complete technical work for SaaS pilot												
Conduct SaaS pilots in conjunction with Lyrasis												
Lead shared sprints in Hydra community to further modularization												
Participate in IIIF A/V spec development												
Conduct initial planning for creation of rightsstatments.org guidance materials												
Develop requirements for Avalon Archivematica integration												
Begin implementation of Archivematica integration in conjunction with Artefactual												
Begin transition of SaaS pilot to operational service in conjunction with Lyrasis												
Carry out second set of shared sprints in Hydra community to further modularization												
Implement preliminary support for the IIIF A/V spec in Avalon and proof-of-concept IIIF A/V client												
Implement functional enhancements and bug fixes												

Activities Year 2	July 2018	Aug. 2018	Sept. 2018	Oct. 2018	Nov. 2018	Dec. 2018	Jan. 2019	Feb. 2019	March 2019	April 2019	May 2019	June 2019
Complete implementation of Archivematica integration in conjunction with Artefactual												
Complete transition of SaaS pilot to operational service in conjunction with Lyrasis												
Continue second set of shared sprints in Hydra community to further modularization												
Complete preliminary support for the IIIF A/V spec in Avalon and a proof-of-concept IIIF AV client												
Implement rightsstatement.org support and test guidance materials												
Wrap up code modularization, IIIF A/V, rightstatement.org work, and SaaS operational service details												
Write final documentation												
Define and implement structure for ongoing Avalon project management and development												
Implement functional enhancements and bug fixes												

DIGITAL PRODUCT FORM

Introduction
The Institute of Museum and Library Services (IMLS) is committed to expanding public access to federally funded digital products (i.e., digital content, resources, assets, software, and datasets). The products you create with IMLS funding require careful stewardship to protect and enhance their value, and they should be freely and readily available for use and re-use by libraries, archives, museums, and the public. However, applying these principles to the development and management of digital products can be challenging. Because technology is dynamic and because we do not want to inhibit innovation, we do not want to prescribe set standards and practices that could become quickly outdated. Instead, we ask that you answer questions that address specific aspects of creating and managing digital products. Like all components of your IMLS application, your answers will be used by IMLS staff and by expert peer reviewers to evaluate your application, and they will be important in determining whether your project will be funded.

Instructions
You must provide answers to the questions in Part I. In addition, you must also complete at least one of the subsequent sections. If you intend to create or collect digital content, resources, or assets, complete Part II. If you intend to develop software, complete Part III. If you intend to create a dataset, complete Part IV.

PART I: Intellectual Property Rights and Permissions

A.1 What will be the intellectual property status of the digital products (content, resources, assets, software, or datasets) you intend to create? Who will hold the copyright(s)? How will you explain property rights and permissions to potential users (for example, by assigning a non-restrictive license such as BSD, GNU, MIT, or Creative Commons to the product)? Explain and justify your licensing selections.

The authors and/or their institutions will hold the copyright of reports and other documentation. We will assign the Attribution 4.0 International Creative Commons license to all outputs (reports, papers, etc.) so that the wider community can use it. http://creativecommons.org/licenses/by/4.0/

Copyright to software developed as part of the project will be held by the individual developers and/or their institutions but will be released as open source under the Apache 2.0 license currently used by Avalon and by the Hydra project.
https://www.apache.org/licenses/LICENSE-2.0
https://wiki.duraspace.org/display/hydra/Hydra+Project+Intellectual+Property+Licensing+and+Ownership

A.2 What ownership rights will your organization assert over the new digital products and what conditions will you impose on access and use? Explain and justify any terms of access and conditions of use and detail how you will notify potential users about relevant terms or conditions.

The final report will be ingested into Northwestern University's Institutional Repository and will be publicly available. The code created will be available through GitHub or similar code sharing site.

A.3 If you will create any products that may involve privacy concerns, require obtaining permissions or rights, or raise any cultural sensitivities, describe the issues and how you plan to address them.

N/A

Part II: Projects Creating or Collecting Digital Content, Resources, or Assets

A. Creating or Collecting New Digital Content, Resources, or Assets

A.1 Describe the digital content, resources, or assets you will create or collect, the quantities of each type, and format you will use.

N/A

A.2 List the equipment, software, and supplies that you will use to create the content, resources, or assets, or the name of the service provider that will perform the work.

N/A

A.3 List all the digital file formats (e.g., XML, TIFF, MPEG) you plan to use, along with the relevant information about the appropriate quality standards (e.g., resolution, sampling rate, or pixel dimensions).

N/A

B. Workflow and Asset Maintenance/Preservation

B.1 Describe your quality control plan (i.e., how you will monitor and evaluate your workflow and products).

N/A

B.2 Describe your plan for preserving and maintaining digital assets during and after the award period of performance. Your plan may address storage systems, shared repositories, technical documentation, migration planning, and commitment of organizational funding for these purposes. Please note: You may charge the federal award before closeout for the costs of publication or sharing of research results if the costs are not incurred during the period of performance of the federal award (see 2 C.F.R. § 200.461).

N/A

C. Metadata

C.1 Describe how you will produce any and all technical, descriptive, administrative, or preservation metadata. Specify which standards you will use for the metadata structure (e.g., MARC, Dublin Core, Encoded Archival Description, PBCore, PREMIS) and metadata content (e.g., thesauri).

N/A

C.2 Explain your strategy for preserving and maintaining metadata created or collected during and after the award period of performance.

N/A

C.3 Explain what metadata sharing and/or other strategies you will use to facilitate widespread discovery and use of the digital content, resources, or assets created during your project (e.g., an API [Application Programming Interface], contributions to a digital platform, or other ways you might enable batch queries and retrieval of metadata).

N/A

D. Access and Use

D.1 Describe how you will make the digital content, resources, or assets available to the public. Include details such as the delivery strategy (e.g., openly available online, available to specified audiences) and underlying hardware/software platforms and infrastructure (e.g., specific digital repository software or leased services, accessibility via standard web browsers, requirements for special software tools in order to use the content).

N/A

D.2 Provide the name(s) and URL(s) (Uniform Resource Locator) for any examples of previous digital content, resources,

or assets your organization has created.

N/A

Part III. Projects Developing Software

A. General Information

A.1 Describe the software you intend to create, including a summary of the major functions it will perform and the intended primary audience(s) it will serve.

Avalon is an open source software package for managing and providing access to large collections of audio and video materials. Planned, designed, and built with support from prior grants from IMLS and the Andrew W. Mellon Foundation, Avalon provides to library, archives, and cultural heritage institutions what existing commercial, open source media systems, and digital repository platforms do not adequately address: an affordable, community driven, preservation ready AV delivery system.
This existing software package will be integrated with the existing Hydra stack and at the same time will be modified in its existing form to work in a cloud based delivery environment.
The other two goals are to connect Avalon to media preservation systems and work-flows and to enable interoperability with existing and emerging scholarly tools and protocols.
These activities will improve user experiences by offering the ability to embed AV resources within larger digital collections, to give developers and institutions the opportunity to decide how they want to deliver their content, either with a locally run system or a cloud based system. It will allow archivists and digital preservation specialists the tools they need to perform digital preservation activities on their AV materials and will allow researchers more flexibility when working with digital content.

A.2 List other existing software that wholly or partially performs the same functions, and explain how the software you intend to create is different, and justify why those differences are significant and necessary.

There are a few AV streaming services used in higher education to deliver AV materials, but none of them offer the content management capabilities of Avalon as it currently exists. With this grant, we would expand what we are offering to include preservation services, which none of the existing vended solutions offers.
Hydra, Islandora, DSpace, and ContentDM all offer various types of content management capabilities, but none work well with AV materials and none of them currently offer the ability to deliver AV content directly within their systems.
What we are proposing with this grant is to integrate Avalon into the existing Hydra stack to provide AV content directly in a well-supported, Open Source system while also introducing digital preservation activities to the system.

B. Technical Information

B.1 List the programming languages, platforms, software, or other applications you will use to create your software and explain why you chose them.

Ruby on Rails
Fedora Repository
Hydra
Red5 Media Server
Adobe Flash Media Server
Wowza Media Server
Amazon Web Services

Avalon is a pre-existing software built on the Open Source tools, Fedora and Hydra. Hydra is built using Ruby on Rails. Red5, Adobe Flash Media Server, and Wowza are both well used and supported tools in the AV delivery sphere and are the gold standard tools for delivery and streaming. Amazon Web Services is well used within the Library/Archives community and is the basis for DuraCloud and a few dark-archival systems. All of these tools were selected because of the wide adoption in higher education and with cultural institutions.

B.2 Describe how the software you intend to create will extend or interoperate with relevant existing software.

As explained above, one of the major outcomes of this work is to integrate Avalon into the larger Hydra stack.
In addition, the digital preservation work, would allow easier flow of materials marked for digital preservation into existing digital preservation systems like APTrust, Chronopolis, and DPN. With the implementation of a IIIF AV protocol, better interoperability would exist for content, despite the platform it resides on.

B.3 Describe any underlying additional software or system dependencies necessary to run the software you intend to create.

There will be three ways to run Avalon at the end of this grant.
In its current form, as a standalone software system with the dependencies currently listed on the Avalon website - http://www.avalonmediasystem.org/software
It will be offered as a series of gems within the larger Hydra framework, which would require all of the necessary components to run Hydra, including Fedora 4.0.
The SaaS offering would not require additional software for the institutions implementing it as their solution.

B.4 Describe the processes you will use for development, documentation, and for maintaining and updating documentation for users of the software.

The team currently uses GitHub to document development activities and we have updated technical information on the Avalon wiki. We have also been posting videos of sprint progress and will continue to do so.

B.5 Provide the name(s) and URL(s) for examples of any previous software your organization has created.

http://www.avalonmediasystem.org/
https://images.northwestern.edu/
NOTIS - https://en.wikipedia.org/wiki/NOTIS
http://sites.northwestern.edu/nubooks/

C. Access and Use

C.1 We expect applicants seeking federal funds for software to develop and release these products under open-source licenses to maximize access and promote reuse. What ownership rights will your organization assert over the software you intend to create, and what conditions will you impose on its access and use? Identify and explain the license under which you will release source code for the software you develop (e.g., BSD, GNU, or MIT software licenses). Explain and justify any prohibitive terms or conditions of use or access and detail how you will notify potential users about relevant terms and conditions.

All software is currently released, and will continue to be released using an Apache 2.0 Open Source License, which allows nearly all forms of commercial and noncommercial reuse and creation of

derivative works. This is consistent with the Project IP Licensing and Ownership framework. License terms will be clearly presented via README and/or LICENSE files distributed as part of the product and made available on GitHub along with the source code. We are not imposing any ownership or access rights.

C.2 Describe how you will make the software and source code available to the public and/or its intended users.

We currently make this product available through GitHub, and offer a testing service through the Avalon Media System website. These will continue.

In addition, the software will be available through existing Library vendors. Currently Data Curation Experts offers limited for-fee Avalon setup and hosting services, and with the results of the SaaS pilot, Lyrasis and/or DuraSpace may also offer these services.

C.3 Identify where you will deposit the source code for the software you intend to develop:

Name of publicly accessible source code repository: GitHub

URL: https://github.com/avalonmediasystem/avalon

Part IV: Projects Creating Datasets

A.1 Identify the type of data you plan to collect or generate, and the purpose or intended use to which you expect it to be put. Describe the method(s) you will use and the approximate dates or intervals at which you will collect or generate it.

N/A

A.2 Does the proposed data collection or research activity require approval by any internal review panel or institutional review board (IRB)? If so, has the proposed research activity been approved? If not, what is your plan for securing approval?

N/A

A.3 Will you collect any personally identifiable information (PII), confidential information (e.g., trade secrets), or proprietary information? If so, detail the specific steps you will take to protect such information while you prepare the data files for public release (e.g., data anonymization, data suppression PII, or synthetic data).

N/A

A.4 If you will collect additional documentation, such as consent agreements, along with the data, describe plans for preserving the documentation and ensuring that its relationship to the collected data is maintained.

N/A

A.5 What methods will you use to collect or generate the data? Provide details about any technical requirements or dependencies that would be necessary for understanding, retrieving, displaying, or processing the dataset(s).

N/A

A.6 What documentation (e.g., data documentation, codebooks) will you capture or create along with the dataset(s)? Where will the documentation be stored and in what format(s)? How will you permanently associate and manage the documentation with the dataset(s) it describes?

N/A

A.7 What is your plan for archiving, managing, and disseminating data after the completion of the award-funded project?

N/A

A.8 Identify where you will deposit the dataset(s):

Name of repository:

URL:

A.9 When and how frequently will you review this data management plan? How will the implementation be monitored?

N/A