



NATIONAL ENDOWMENT FOR THE

Humanities

DIVISION OF PRESERVATION AND ACCESS

Narrative Section of a Successful Application

The attached document contains the grant narrative of a previously funded grant application. It is not intended to serve as a model, but to give you a sense of how a successful application may be crafted. Every successful application is different, and each applicant is urged to prepare a proposal that reflects its unique project and aspirations. Prospective applicants should consult the NEH Division of Preservation and Access application guidelines at <http://www.neh.gov/divisions/preservation> for instructions. Applicants are also strongly encouraged to consult with the NEH Division of Preservation and Access staff well before a grant deadline.

Note: The attachment only contains the grant narrative, not the entire funded application. In addition, certain portions may have been redacted to protect the privacy interests of an individual and/or to protect confidential commercial and financial information and/or to protect copyrighted materials.

Project Title: Quality Control Tools for Video Preservation

Institution: Bay Area Video Coalition (San Francisco, CA)

Project Director: Moriah Ulinskas

Grant Program: Research and Development

Attachment 3: Narrative

Bay Area Video Coalition – “Quality Control Tools for Video Preservation”

A. Significance

Need for Quality Control Tools

Currently, extraordinary amounts of time are required to: 1) Transfer analog tape recordings to digital file formats while meticulous documentation and adjustments are made; 2) Review, in-full, potentially problematic files, identifying any compromises in tape playback or digitization tools; 3) Analyze and identify problems and their sources; and 4) Develop potential solutions to tape degradation, test them, and apply them. Video preservation through digitization is an activity that is complex but essential for all types of heritage organizations to begin. For archivists digitizing video, the diversity and complexity of possible solutions can introduce issues that may go unnoticed, leading to the risk of handing forward a historical defective record. BAVC’s work with diverse archives and partner institutions has identified this as a serious national issue.

The lack of industry-standard quality control tools and expertise in the audiovisual preservation field results in technical compromises that can go undetected, with minor flaws in a digitization station undermining all resulting digital media. Given the increasing urgency to digitize analog audiovisual collections of all kinds, a quality control toolkit for archivists is imperative to dissect, analyze and explain the resulting signals and file structure produced by various digitization workflows.

Archivists require a means to assess their digitization work to demonstrate that their procedures are trustworthy, identify compromises in their setup, or receive facilitated guidance and prioritization of quality control efforts. Existing tools that flag digitization errors are rare, expensive and not comprehensive. They are designed for broadcast television standards and regulations, not for video preservation purposes, nor are they relevant for small humanities archives containing documentation tapes or mixed collections. There is a strong need across the field, therefore, for shareable, open source quality-control tools in order to increase the affordability and practicality of high-quality digitization.

Michael Angeletti, Moving Image Digitization Specialist at Stanford University Libraries, the department that reformats for preservation a wide variety of humanities-based audiovisual collections for Stanford, reports that:

...I would look for integrated quality assurance tools that would verify that each file is well-formed and meets a customizable quality threshold, for example the average RF levels during playback and signal clipping. Because many tools currently available are built for use in post-production workflows, the analysis and reporting functions are limited.

In order to develop new software tools, BAVC proposes continuing its partnership with independent consultant Dave Rice and the Dance Heritage Coalition (DHC) to use the current Secure Media Network repository (SMN - described in History section below), as a testing ground to develop the *Quality Control Tools for Video Preservation* for handling major risk

points in the process, the transfer of legacy assets to repositories, and the capture of metadata standards to benefit the fields of preservation and humanities.

Broad Application with Free Adoption

Audiovisual materials are an ever-increasing part of the archival and cultural record. Digital tape formats for standard-definition video were introduced in the early 1990s and are no longer a viable option for the long-term preservation of analog video. To avoid degradation and obsolescence, existing analog and tape-based video need to be migrated into and managed within a digital environment. This digitized legacy video combined with the exponentially increasing amount of easy-to-create 'born-digital' video will soon overwhelm collecting institutions and complicate the IT infrastructure of small-to-medium archives. Many institutions are in the double-bind of losing access because of 1) format obsolescence if material is not migrated; and 2) inability to access those digital files if there is no reliable way to manage them. The innovations and solutions developed in this project will benefit a wide array of video image producers, users, preservationists and archivists. BAVC and its partners are dedicated to creating open source, non-proprietary utilities that can be freely adopted across fields, and to incorporating archival standards and best practices for reformatting and capturing metadata that enables the long-term preservation and of and access to the original artifact, the digital object, and the associated catalog record.

Cost Reduction and Reliability through Automation

Existing tools to achieve quality control of video streams tend to be expensive and out of the reach of many BAVC partners and other small archives. Available tools such as Final Cut Pro, Quicktime and Compressor are designed for the needs (and budgets) of video production and post-production rather than long-term maintenance and preservation. BAVC and the DHC have been pioneers in the development and testing of much-needed open source alternatives for the quality control of legacy archival audiovisual materials. The partners have worked on prototypes for the *Quality Control Tools for Video Preservation*, the implementation of open source repository software Archivematica, and the inclusion of open source reporting tools that use ffprobe and MediaInfo.

An integral part of this project's work is to establish an automated processes for mapping and translating existing database records to a single metadata standard, create new catalog records, and extract embedded metadata from digital files. By limiting the amount of human time required to process records, these newly automated processes will ensure greater control and integrity throughout the records while limiting input error and substantially reducing cost. Cost reduction through the exploitation of computing functionality is a major task identified by the NSF/LOC Digital Archiving workshop as a requirement for encouraging the adoption of digital preservation activities. The development of a similarly low-cost alternative for large files is, therefore, critical.

Access for Study

Humanities research is made possible by access to original source materials. Advances in technology have resulted in increased accessibility to texts and still images, but access to moving images lags far behind still images in quantity and quality. Even if a researcher were able to travel to an archive holding audiovisual materials, the content may still be inaccessible due to

condition, technology obsolescence, or lack of proper playback device. *Quality Control Tools for Video Preservation* is a crucial next step to decrease cost, increase speed and accuracy of turnaround, and provide a model for preservation-level standards and best practices built on open source and cost-effective utilities. The proposed areas for research and development detailed in this proposal will be a huge step forward in creating a model for access and maintenance of digital audiovisual assets for the long-term.

BAVC has identified these quality control tools as a field-wide need, and so we will be developing a website that will be a resource for accessing and distributing not only the tools, but also information about tool development, as well as case studies for testing. The website will familiarize and educate users with these tools, assisting in analyzing the resulting data. The website is an ongoing part of the project; it will provide a vehicle for distributing videos, papers and other documentation that will aid archivists and digitization technicians in using the *QC Tools* via the website.

The tools developed for the project will also be available on GitHub, Inc. GitHub is in use by the software and open source computing community. It is self-described as “the best place to share code with friends, co-workers, classmates, and complete strangers. Over a million people use GitHub to build amazing things together.” GitHub will be an invaluable tool for collaboration between BAVC and its testing-ground collaborators, and documenting the results of these implementation tests.

As an example of how the *QC Tools* will be adopted and updated, BAVC currently works with the San Francisco Poetry Center at San Francisco State University (SFSU). Their Digital Information Virtual Archive, (<https://diva.sfsu.edu/collections/poetrycenter>) (DIVA) chronicles the history of American poets, particularly in the 1960s and 1970s. In addition to the Poetry Center’s archival collections, DIVA houses other archives such as the Bay Area News Archive and the College of Ethnic Studies (and others). Video works from this collection are used by students and scholars throughout the world. Currently BAVC works with the Poetry Center archivist and director to prioritize recordings for preservation and make these works available through the DIVA website. The administrator for the DIVA site is one of the participating partners in this research project and will be trained in the *QC Tools* as they are developed, giving the development team feedback on the tools’ usability, and making ongoing recommendations for the tools’ broad dissemination. This is one of many similar partnerships BAVC has with humanities-related archives nationally. We see this partnership as widely replicable across academic areas, and believe it will make hundreds of currently inaccessible video resources available for study.

B. Background of Applicant

The Bay Area Video Coalition (BAVC) is a nonprofit media arts center founded in 1976 by a coalition of media makers and activists who wanted to find alternative, civic-minded applications for emerging video technologies. BAVC’s continuing mission is to inspire social change by enabling the sharing of diverse stories through art, education, and technology. The preservation of audiovisual content to ensure access to significant cultural artifacts for future generations is integral to BAVC fulfilling its mission. BAVC is one of the nation’s longest-standing nonprofit organizations dedicated to video and audio preservation. Established in 1996 in partnership with

the Getty Research Center, BAVC's Preservation Department developed practices and services that helped to establish the field of video preservation and for the last 16 years has provided critical preservation services to artists and cultural institutions of all sizes.

BAVC's Preservation Department works with museums, artists and cultural institutions around the world to remaster, transfer and archive seminal creative and historical works from video and audio tape. BAVC's preservation team has provided services for distinguished organizations such as the The Kitchen, Video Data Bank in Chicago, SF MOMA, BAM/PFA, Walker Arts Center, U.C. Berkeley, U.C. Davis, and San Francisco Ballet among others.

BAVC has developed long-term, trusted relationships within the humanities and independent media fields, providing the history and expertise to perform this technical work to advance the field. BAVC consistently employs developing technologies and best practices, and shares these practices openly and widely among our peer and partner organizations. BAVC preservationists have also trained other culturally-focused organizations, such as the Getty Center, in developing their video preservation systems.

BAVC makes it a priority to promote best preservation practices by presenting at the most important well-attended preservation conferences, including the Association of Moving Image Archivists Conference, The American Institute for Conservation of Historic and Artistic Works conference and the Andrew W. Mellon Symposium on Technical Issues of Time Based Media. BAVC also publishes scholarly work in publications such as Electronic Media Review.

In addition, BAVC develops training materials for independent video producers, making BAVC uniquely positioned to develop localized preservation packages and provide targeted training to key stakeholders on the use of the proposed *Quality Control Tools for Video Preservation*.

Project Partners: Dave Rice and the Dance Heritage Coalition (DHC)

Independent consultant Dave Rice has been a key member of BAVC's Preservation Program since 2008. Rice designs tools and workflows for processing media and metadata in efficient and scalable ways. Rice worked as an archivist or consultant to organizations such as the City University of New York, Democracy Now!, WNET, United Nations, Downtown Community Television, and WITNESS. His work focuses on the utilization and development of open source software to facilitate audiovisual preservation objectives with particular emphasis on quality control analytics and metadata utilization. Rice has managed several research and development projects focuses at addressing issues within the audiovisual preservation community through open source software development, information aesthetics, and technical metadata analysis.

Rice brings to the project three collaborating partners: Jerome Martinez, the lead developer of MediaInfo, an open source audiovisual assessment tool. A frequent collaborator with Rice, Jerome is an expert in digital file inspection, quality control, software design, and technical metadata. Mark Heath, an independent programmer with expertise in video signal and data stream analysis. Heath brings to the project the video engineering expertise and skills to enable audiovisual signal analysis in a manner that is efficient in time and processing resources. Devon Landes, an archivist, specializing in broadcast and advertising collections. Landes brings project

management and technical writing expertise as well as a perspective from the rigorous quality control environments of broadcast television and advertising.

The Dance Heritage Coalition (DHC) is a national alliance of institutions holding significant collections of materials documenting the history of dance. The DHC was founded in 1992 to address problems in documenting and preserving dance, which were identified in a field-wide study commissioned by The Andrew W. Mellon Foundation and the National Endowment for the Arts. Coalition members include American Dance Festival, Arizona State University, Dance Notation Bureau, Harvard University, Jacob's Pillow Dance Festival, The New York Public Library for the Performing Arts, Library of Congress, Museum of Performance & Design (San Francisco), The Ohio State University, and the University of California, Los Angeles. The DHC's well-established dedication to developing innovative technological and collaborative approaches to fulfilling its mission for preservation and access and its generosity in sharing its work products make the DHC an ideal test bed for research and development in digital video repository management, the focus of this project proposal.

With operating funds from the Mellon Foundation and additional project support from the NEH, IMLS, and NEA, the Delmas Foundation, the Hewlett Foundation, the Doris Duke Charitable Trust, and others, the DHC has identified, developed, and implemented national projects to improve the preservation and accessibility of materials documenting the cultural and aesthetic history of dance in the United States. The DHC focuses especially on three areas: 1) the continuing documentation of dance employing both traditional methods and developing technologies; 2) the preservation of existing documentation; and 3) strategies to ensure enduring access to dance legacy materials. The DHC also disseminates its methods, standards, and resources to its constituency as well as to related groups and individuals who can put these products to their own uses.

C. History, scope, and duration

History

Quality Control Tools for Video Preservation proposes to make use of data gathered through BAVC's extensive 2007-2012 preservation project for the Dance Heritage Coalition, and the resultant Secure Media Network (SMN) on which the DHC media archive is now stored at BAVC. The goal of the project is simple: to cut down the time and expense it takes to perform high-quality video preservation, enabling preservationists to direct time towards issues that are *solvable*. For example, identifying tapes that would benefit from a second transfer, saving not only the precious time of preservationists and institutional resources, but also giving collections a necessary advantage in the bigger race against time to preserve their significant cultural artifacts.

Role of Secure Media Network

Providing tremendously valuable data and a model for collaboration for the development of these new preservation tools, the SMN has served as a union catalog and secure digital repository for the diverse collections of the dance community. The participating institutions individually manage parts of America's dance heritage through very different approaches to metadata and media management. The SMN incorporates metadata translators to allow data from the partners to be managed through a selected conformance metadata standards within the SMN repository, enabling faceted browsing, search, and analysis of data across the partnership.

The SMN also hosts and manages the digital results of the Dance Heritage Coalition's audiovisual preservation efforts. BAVC and partners have adapted traditional repository systems and microservices to serve the large file sizes and complex architectures involved in preservation-quality digital video. The development of the SMN has helped clarify objectives, served as a model and provided recommendations of video digitization within the archival community. As a result, BAVC and partners have established tools to enact the workflow and strategies outlined by the Open Archival Information System (OAIS) model.

Although the SMN made significant progress in utilizing the OAIS standard in digital video preservation by adapting open source software, establishing packaging standards for digitization work, and designing assessment tools, most of these tools are not properly designed for the large files, quality control issues, or technical responsibilities of video preservation. *Quality Control Tools for Video Preservation* is the next phase in this important preservation effort to create tools designed for batches of large audiovisual files to better serve small to medium sized cultural institutions and their legacy collections.

Digitization Stations

Concurrent with the development of the SMN, BAVC built prototypical digitization stations that will provide additional testing opportunities for the new quality control tools. Under the guidance of BAVC staff, digitizing technicians at these stations are trained in preservation best practices, signal chain workflow and metadata standards:

- The Museum of Performance and Design digitizing station was designed and built by BAVC in 2010. This station, which is designed to focus on analog materials most common in the institutions served by the MPD, was housed in the MPD for two years during which time two fellows and the MPD archivist were trained in signal workflow and preservation digitization practices.
- The second station, at the Dance Notation Bureau, was designed and installed by BAVC and Dave Rice in June 2011 and focuses on a mix of analog and digital-tape based recordings. There has been one fellow trained on signal flow and preservation practices on this station, who remains there today.
- A third station is slated to be installed in Washington D.C. in the Summer of 2012, with plans to add at least one more, in Austin, TX, in 2013.

These stations are highly controlled environments with well-documented processes, making them a perfect case study site for *Quality Control Tools for Video Preservation*. At this phase, more than 100 legacy video assets have been digitized and submitted to the SMN repository, and 27,000 catalog records of video assets have been assembled in a publicly available and searchable online catalogue and digital repository (<http://archive.danceheritage.org>), and many thousands more are expected from the other members in months and years to come.

Audio/Visual Artifact Atlas (AVAA)

The goal of *Quality Control Tools for Video Preservation* is to create new software tools that have the potential to benefit all areas of the humanities. A crucial step in this effort was the creation of an accessible error database. This database, created during the development of the SMN, is now aggregated and available to the public in the Audio/Visual Artifact Atlas (AVAA).

Archivists can now use the AVAA material to learn about technical issues and apply the information to the quality control process to better inspect the video products. Similarly, entities working on reformatting may refer to the AVAA as an independent resource in order to identify and name issues encountered while transferring material. The AVAA provides a common vocabulary for archivists, the digitization service providers with whom they work and other users of archival content, and offers supporting examples for discussing such issues with one another.

BAVC hosts the AVAA, but Stanford University and NYU are significant contributors to this important prototype project. The AVAA has received a strong initial response from the preservation community. It represents a promising new tool that serves the front end in a series of steps which comprise this larger *Quality Control Tools for Video Preservation* project - that will both use the data on the AVAA and add to it.

Quality Control Tools for Video Preservation will be founded upon the AVAA and all the important work accomplished in partnership with the DHC, and will enable BAVC and partners to develop the next phase of software tools to benefit all areas of the humanities dealing with the digitization of their legacy works. This project will result not only in the software tools, presentations and white papers, but in the education and training of important partners in institutions around the country on how to apply the tools to the digitization of their archives.

Presentations and Publications

BAVC's senior preservation staff and team members present research and serve on panels at regional and national preservation conferences, including the Association of Moving Image Archivists (AMIA) annual symposia, and consult with a range of preservation stakeholders, from newcomers to leading experts in the field. Preservation Director Moriah Ulinskas serves as the Chair of the Diversity Committee for AMIA, while Preservation Specialist Lauren Sorensen serves as the Co-Chair for the Independent Media Committee. Additionally, Dave Rice serves as the Chair of the Open Source Committee for AMIA. Relevant presentations and publications include:

- 1996 "PLAYBACK: A Preservation Primer for Video" published in conjunction with a conference produced by BAVC and funded by the Getty Grant program featuring white papers by leading preservationists from Rochester Institute of Technology, Stanford University Libraries, American Institute for Conservation, and the National Archives.
- 2000 "TechArchaeology" a symposium on preservation of time-based art and installation art, sponsored by BAVC, bringing together conservators, curators and technical experts around the "Seeing Time" exhibition at SFMOMA.
- 2003 "PLAYBACK: Preserving Analog Video" DVD published, documenting the step-by-step conversation and preservation of Ant Farm's seminal work "The Eternal Frame" in partnership with the Tate Modern
- In November 2009, Dave Rice, Angelo Sacerdote and Skip Elsheimer presented "Video Digitization Workflows and Challenges". The presentation culled an extensive set of digitized media in order to demonstrate quality control issues, categorize them, and propose responses. This presentation along with audience feedback led to early planning for this proposal.

- In 2010, Mark Hellar presented, “Virtualizing *Agent Ruby*: Collecting Web Art,” with Jill Sterret at DOCAM (Documentation and Conservation of the Media Arts Heritage) Conference.
- In April 2010, Mark Hellar presented, “*The Virtualization of Agent Ruby: A Case Study in the Preservation of Web Art*” at Harvard University’s Andrew W. Mellon Symposium in Conservation Science: Technical Conservation Issues of Time-Based Media.
- 2010 BAVC’s “Preserving Video: The Digital Transition” presented at the Andrew W. Mellon Symposium on Technical Issues of Time Based Media.
- 2011 Moriah Ulinskas presented “Documentation, Preservation and Access in the Age of Digital Media and Information” to the Arts Loan Fund at the Northern California Grantmakers.
- 2011 Lauren Sorensen and Dave Rice presented “Secure Media Network: Building a Digital Repository for a Diverse Coalition of Analog Video Collections” at the Association of Moving Image Archivists’ Conference.
- In November 2011, Lauren Sorensen presented as a part of the panel “Recommendation 4.2: Orphan Video Preservation as Regional and Federal Concern” at the Association of Moving Image Archivists Conference.
- On March 15, 2011 Dave Rice and Skip Elsheimer (technical advisor to this project) presented “Transcoding for Access and Preservation” at the Screening the Future conference at the opening of PrestoCentre. The presentation reviewed how to control processing of audiovisual data to minimize artifacts and to increase accessibility.
- 2011 Moriah Ulinskas presented “Forward to the Past” at the Grantmakers in the Arts Conference.
- April 2012 “Codec Comparison for the Digital Preservation of Analog Video” by Angelo Sacerdote and Lauren Sorensen published in the Electronic Media Review, by Electronic Media Group of the American Institute for Conservation of Historic and Artistic Works.
- April 27, 2012 Lauren Sorensen presented “Business of Audiovisual Preservation” panel and the Society of California Archivists.
- May 11, 2012 Lauren Sorensen presented “Capture Software for the Preservation of Analog Video” at The American Institute for Conservation of Historic and Artistic Works conference in May 2012.

Scope of project

A variety of humanities-related archives have expressed a strong interest in participating in the piloting of *Quality Control Tools for Video Preservation*, including the San Francisco Public Library, the Pacific Film Archive, the San Francisco Exploratorium, the NYU Libraries Preservation Department and the Outfest/Legacy Project at UCLA among others. Please see list below of individuals committed to participating in the pilot.

The Library of Congress notes in their guidelines for Sustainability of Digital Formats:

[Seven sustainability] factors influence the likely feasibility and cost of preserving the information content in the face of future change in the technological environment in which users and archiving institutions operate. They are significant whatever strategy is adopted as the basis for future preservation actions: migration to new formats, emulation of current software on future computers, or a hybrid approach.

These factors include: disclosure, adoption, transparency, self-documentation, external dependencies, impact of patents, and technical protection mechanisms. *QC Tools* will aid archivists analyzing these factors and making informed choices for sustainability, understanding how media changes through the process of digitization, and trusting that their digital preservation work is accurate, well-documented, and transparent.

QC Tools will benefit not only moving image archives but paper archives that include video collections that may fall into neglect. The rapid rate of deterioration makes the preservation of magnetic media-based archival elements of foremost import when deciding on preservation strategies; however, because the knowledge base is so specialized, paper and photographic archivists lack a way to evaluate digitization measures. *QC Tools* will make it easier and more approachable for these archives to preserve and maintain digitized moving image materials.

Duration

The initial phases of the project began in 2007. From January 2013-January 2015, we propose to develop, pilot and disseminate the *Quality Control Tools for Video Preservation* with support from the National Endowment for the Humanities. BAVC will maintain and further develop the tools to ensure their relevancy to the field. These open source tools have the potential to benefit the field for years to come.

Though codecs change as new video systems are created by the video-making community, the openly-licensed libraries *libavcodec* and *libavformat* have a broad base of support and are continually being developed by the computing community, and so make our suite of tools, which will apply to these codecs, very sustainable for the long term. BAVC and partners are aware of complex issues around versioning control, operating system compatibility, and other variables beyond our control. However, we are confident that the wide application of codecs and the community behind open source applications for use in the archival community is strong enough for this tool to have a measurable effect for the long-term. Through disseminating the *QC Tools* on the website, to the partner and pilot institutions as well as at conferences to specialists and non-specialist audiences in the archive community, we can encourage and promote use and development well beyond the NEH grant period.

D. Methodology and Standards

The design of the tools and documentation within this project will strongly encourage best practices and standards in audiovisual digitization of analog material for users. These standards, culled from procedures of the Federal Agencies Digitization Guidelines Initiative, PrestoCentre, and Stanford University include:

- the utilization of uncompressed or lossless codecs and avoidance of any intermediate lossy signal processing,
- adherence to chroma subsampling, frame sizes, frame rates, and audio channel configurations that correspond to the specifications of the analog source material,
- use of bit depths responsive to the full detail and signal of the analog source, chiefly 10 bit for video sampling and 24 bit for audio sampling,
- choice of codecs and containers that are non-proprietary, openly-documented and supported by sustainable communities,

- avoidance of using specifications or features within codecs or containers that add unnecessary complexity to resulting digital preservation files.

Although these guidelines reflect our team's recommendations for the digitization of analog video material for preservation, many collections are currently digitized without regard to standards due to utilization of lossy codecs and tampering with technical specifications of the source signal (such as frame size, chroma subsampling and frame rate). Our tools will work for these types of collections as well: we will incorporate the openly-licensed libraries *libavcodec* and *libavformat* to enable our tools to decode and perform analysis on nearly any video file regardless of the circumstances of its digitization. Thus the *QC Tools* will not be limited to a small subset of audiovisual specifications, but can be applied to a variety of formats including proprietary ones. The software will test for the preservation of the analog signal throughout the digitization process, and alert the user to quantified discrepancies caused by utilization of lossy codecs.

Regarding metadata standardization, the tools will analyze a set of videos to produce data files that documents technical aspects, its contained tracks, as well as the packets or frames of audiovisual data within those tracks. While this information will be displayed to the user through graphs, charts, visual comparisons, and summarized reports, the underlying quality control data will also be stored.

Since most digital audiovisual quality control tools that do exist are proprietary, there are significant differences in the scope, structure, and format from one tool to the next. Vendor quality reports and methods for storing data are very different, posing significant challenges in interoperability. For this project, BAVC will design a data format for storing assessed data from the analysis: rather than generating yet another proprietary solution, our team will review existing metadata standards and data formats that track technical details at the levels of tracks, packets and frames in order to author an open specification and metadata standard for storing complex quality control data.

BAVC's design and documentation of an open quality control technical metadata format will enable more opportunities for interoperability so that data gathered through the analysis of these tools can be utilized in other environments or imported into local database. Releasing a specification of this quality control metadata will enable for the use of tools outside this project. It will be possible for data from a proprietary system to be translated to the new quality control metadata format. As a result, that data can be interpreted and assessed through the new tools, and in turn supplement our open source quality control tools.

The initial work produced through the partnership between BAVC and the DHC had a clear focus on metadata standardization, controlled metadata workflows, and adoption and definition of best practices. Because of the diverse and decentralized nature of the DHC's archives, the partner collections have deployed a broad set of standard and non-standard metadata to document the cumulative collections. The approach has been non-intrusive to the partners: the metadata practices of each contributing institution are defined as a metadata standard, then a translation tool conforms the partner's metadata to the common selected standard, PBCore 1.2.1. Within this approach BAVC has developed MARC to PBCore translators, translation tools from

custom Filemaker and Microsoft Access databases and other standards. The interfaces present the union catalog as PBCore, but also manage links to original submitted metadata records if the translation is not semantically lossless (such as with MARC to PBCore).

Within the digitization work at BAVC, we are employing PREMIS to track preservation, process, and auditing data of archival assets. PBCore is utilized to track technical metadata of the results of digitization, and METS provides the structural framework of all media and metadata for each package as well as providing the link between the library of archival asset packages and the repository.

For software design and coding, this project will have a significant focus on designing interoperable cross-platform software. The results of this project will function in Mac, Windows, or Linux in both graphical user environments and programmatically via command line tools. For coding quality control filters and analyzers we will be using either C or C++, which are two of the most widely used computer programming languages, and also shared by relevant existing tools such as FFmpeg and mediainfo. Our graphical user interface framework will be Qt which excels as a cross-platform user interface framework and incorporates many of the building blocks that our project needs.

For licensing, the project will be released under an open license that will enable anyone to edit or repurpose this code for any other intention. The software design and integration of existing open source tools will be performed in a manner that enables an open license. While the license selection is pending assessment of integrated tools, we are currently considering LGPLv3+ and BSD as licensing models for the project.

By using an open code repository, GitHub, all changes, drafts, and progress in programming and development software will be available for review. Our project intends to not only make the resulting code available but also the full development process so that onlookers and users can understand the change progress and development that formed the final project.

Audio/Visual Artifact Atlas (AVAA – described above) will be crucial in identifying errors and issues within digitization materials, and subsequently writing code to identify, assess, and prompt a response to the issue. The AVAA will continue to grow simultaneously with the *QC Tools*: the deployment of code to track each issue will enable us to plot our progress as we respond to quality control errors in a prioritized order.

BAVC has confirmed nine partners from institutions of varying sizes to pilot the training and adoption of these tools, with an eye on a summative evaluation, presentations and publication of a white paper at the end of the project which will lead to further opportunities for tool development and dissemination. Confirmed participants for the pilot training are:

- Michael Angeletti, Moving Image Digitization Specialist, Stanford Media Lab
- Tom Colley, Collection Manager , Video Data Bank
- Glenn Wharton, Media Conservator
- Jon Worona, Digital Initiatives Manager, San Francisco Public Library
- Alex Cherian, Film Archivist, SF Bay Area TV Archive

- Jon Shibata, Assistant Film Archivist, Pacific Film Archive
- Ruth Falanga, Director, Library and Information Resources, Exploratorium
- Alice Moscoso, NYU Libraries, Preservation Department
- Kristen Pepe, Project Manager, Outfest/ Legacy Project Archives

These participants, through the training, the continual updating on GitHub, and real-time feedback from their institutions once they are actually using the tools, will provide crucial evaluation of the project throughout the two-year cycle and beyond.

E. Work Plan

The proposed project will span two years. The first part will focus on the development of tools for use in the repository workflow and archival digitization environments. The central piece of the tool set is the development of a data format to document and express quality control, analytical, and visual qualitative data across the frames of a digital video, *QC Tools* report. The data format will be designed to be transparent, understandable, plottable, flexible, and efficient to store analytical data on a frame-by-frame basis. Alongside the development of the data format will be the design, testing, and release of an open source, multi-platform tool to generate this report from referenced digital video files as well as tools to analyze and visualize the collected data according to locally defined priorities and thresholds.

The second part of the development will focus on designing a software application to simplify generation of *QC Tools* reports for batches of digitized video. The software applications will enable the archivist to depict the contents of the *QC Tools* report as graphed data, revealing trends, averages, and extremes of the signals stored within the digitized video. The applications can also be used to set thresholds of tolerance for the extent to which the reported digitized video signal can veer from expected trends of best practices (such as video out of a legal range, extended drop-outs, changes in). The *QC Tools* will also facilitate sorting of collections of digitized video according to the presence of digitization issues to enable the archivist to focus quality control and assessment of video on the digitized content most deserving of attention.

An ongoing focus of the project will be the deployment of the full toolset and dissemination of findings and training materials to archivists, media and preservation specialists, as well as all areas of the humanities where access to reliable information on digital preservation standards and practices is not easily accessible.

As mentioned above, BAVC has confirmed nine partners from institutions of varying sizes to pilot the training and adoption of these tools as well as help evaluate them, which will lead to further opportunities for tool development and dissemination.

Timeline Ongoing

- Through the project's quality control process, BAVC's technicians will seek out, identify, and document a diverse library of video samples of digitization and quality control issues.
- Publish project updates and articles to the project website.

- Propose and present at relevant heritage, archival, and technology conferences on project progress and findings.
- Outreach to educational and humanities communities.

January 2013 – February 2013

Summary: During this phase we will establish project communication tools and the initial project website, and convene the Advisory Board. The first major deliverable is the definition of a data format for quality control metadata for digital video covering containers, streams and frames. The data format will be appropriate for storage within a digital repository or to generate reports or graphs.

- Kick-off meeting with BAVC and project personnel and advisors.
- Establishment of a project website on a subdomain of bavc.org.
- Use project website to describe the agenda, timeline, and intent of the project.
- Incorporate relevant presentations and papers from our technical team into the project website.
- Assemble a technical advisory board to review ongoing project work and deliverables.
- Set up communication tools for our advisory board and software development team, including project management sites, email lists, and collaborative and open editing systems.
- Draft of an implementation of the frame-based checksums to serve the needs of long-term preservation.
- Generate an initial draft of a data format for the expression of quality-control and digital signal processing metadata for digital media files, their encoding streams, and their frames.

March 2013 – June 2013

Summary: During the spring we will develop and release an early version of our *QC Tools* as a command line application. This application will be able to assess digital video to generate the defined quality control data format covering key characteristics of the audiovisual signals, such as extremes of signal values per frame, quantified drop-out detection, quantified difference from one frame to the next, black levels, frame-based checksums, and frame borders. We will also document workflows for using the application to enable verification and assessments of lossless encoding.

- Document and code a workflow for generation of frame-based checksums.
- Design, document and distribute test methods for digitization systems that employ lossless encoding to facilitate verification of lossless transcoding within digitization procedures.
- Develop and release uncompressed digital video files within multiple combinations of colorspace, bit depths, and pixel format that contain every possible combination of sample values (to facilitate ongoing testing of lossless compression).
- Draft code to perform digital signal processing to report on minimum, maximum and average video sample values to identify improperly calibrated video digitization.

- Release first project application as a cross-platform *QC Tools* command line application to assess video and report on key frame characteristics to produce the quality control data format.
- Collect, document and publish samples of digitization errors.

July 2013 – September 2013

Summary: The command line application will be extended to incorporate identification of gathered quality control issues pertaining to both the audiovisual signal and container format as well as using frame and field data to quantify the role of the playback device or potential for improvement in certain perceived errors. All features of the command line application will be ported to a Qt-based graphical user interface in order to facilitate broader testing and feedback.

- Revise and publish specifications for the data format to report file-based and frame-based quality-control data that project tools will support.
- Extend digital signal processing code to identify discrepancies between digital video data stream and the physical heads that read the magnetic signal via playback of archival analog tape.
- Collaborate with technical advisory committee to identify and prioritize error detection efforts from the accrued sample library and research.
- Extend digital signal processing code to identify other digitization efforts (i.e. cell phone interference affecting the digitization, Digibeta tape surfaces damage effect on video, etc.).
- Extend the *QC Tools* command line application for further discernment of gathered digitization errors.
- Release first alpha of the graphical user interface (GUI) version of the *QC Tools*.
- Encourage wider testing among the technical advisory board, interested parties, and the public.

October 2013 – December 2013

Summary: By the end of 2013 our project team will be collaboratively working towards producing the first beta release of the *QC Tools*. December will include a feature freeze to focus solely on resolving bugs or issues identified from earlier testing. These months will also focus heavily on documentation both on project websites and technical documentation to be incorporated into the produced software.

- Publish results of frame checksum work, along with standards of the output, associated tools to facilitate the output, and workflow recommendations.
- Using the frame checksumming tools, perform stress tests to determine likelihood of unexpected generation of errors within the transcoding process. Publish results.
- Drafting metadata documentation, documenting quality control microservices, documenting scenarios to integrate *QC Tools* into digital repositories. Release of training materials, control files, and expected outputs.
- With support of technical advisory board, publish the recommended quality implementation of quality-control metadata standard.
- Incorporation of settings within *QC Tools* to enable users to prioritize or rank various quality control issues and set threshold for alerts.

- Adding and refining infographic features of the *QC Tools* in order to present visually quality control issues across collections and timelines from digital signal processing reports and gathered data.

January 2014

Summary: The cross-platform Beta of the *QC Tools* will be released. At this time command line and graphical applications will be available for all major operating systems. BAVC will publicize the release and facilitate outreach to target institutions and collections.

- Public release of a beta of the digitization quality control toolset.
- Documentation of application instructions, training materials, and frequently asked questions.
- Technical forums, bug tracker, and feature request sections will be announced to relevant archival communities.
- Collect comments and feedback from the preservation community.

February 2014 – April 2014

Summary: BAVC will incorporate community feedback into an issue tracker and work with the technical advisory committee in order to prioritize further development of quality control assessment tools, infographic capabilities, and user interface design.

- Review comments and feedback, perform testing.
- Software development focus on improvements to the digitization quality-control toolset that target community feedback.
- Work on user-friendliness, software design and workflow.
- Extension of application reporting capabilities in order to accommodate various collection sizes or desired levels of technical verbosity.

March 2014 – June 2014

Summary: The software development team will establish a release schedule and produce regular updates in order to facilitate a fast feedback loop with the technical advisory committee and early testers and users. While the earlier development focused on digital videos that were sourced from analog materials, this section will produce some support to born digital audiovisual recordings such as HDV tape, XDCam discs and file sets, and DV recordings.

- Updated releases of the *QC Tools*, updated user documentation, and support materials.
- With quality-control toolkit in place, begin development of digitization systems for born-digital non-file-based media (such as DV tape or optical discs).
- Develop workflow, documentation, training materials and standardization of deployable systems.
- Continued development of quality-control toolset in response to general feedback and continued project findings.

March 2014 – August 2014

Summary: With establishment of the AVAA library of video digitization error samples and advice of the technical advisory board, we begin research and development of quality control

tools (detection video head switching, electronic interference, Digibeta playback glitches, low and high peaks in YUV sample values, timecode inconsistencies, and other identified errors).

- Draft related documentation and samples on digitization quality control and error detection.
- Assess existing technical specifications of the digital media within the SMN repository, build and document decoder/demuxer materials in alignment with established project goals of preserve the means to playback digital holdings.

July 2014 – August 2014

Summary: The summer will include the release of version 1 of the *QC Tools* with all associated technical documentation, sample libraries, outreach materials, and documentation of workflow scenarios. With the release, BAVC will focus on system deployment and participating in the technical research that is enabled by community use of quality control tools. We will also submit the software code to open software repositories such as those hosted by GitHub (in other project's we've found this step crucial for extending outreach and sustainability as well as bring a vetting process to the open source licensing). Similarly BAVC will submit patches to the homebrew (<http://mxcl.github.com/homebrew/>) project, a software package manager for Mac computers.

- Full release of the *QC Tools* version 1, including dissemination, outreach and documentation.
- Submit code to Github software repository.
- Facilitate support for the *QC Tools* within homebrew.
- Deploy the *QC Tools* to BAVC and DHC partners, and pilot participants along with documentation and training materials. Train on workflow and collect feedback from users.

September 2014 – November 2014

Summary: We will continue to support and maintain all project-generated software and work to foster incorporate of the resulting code into other relevant projects, such as digital repository tools, video processing utilities, and other archival projects. We will focus on outreach, presentations, and participation in archival communities in order to draw attention to the urgent need to audiovisual digitization as well as:

- Make improvements to the *QC Tools* based upon user feedback and results of quality-control toolset.
- Draft final report of project findings and circulate among technical advisory board for comments.
- Disseminate at conferences and meetings.

December 2014 - January 2015

Summary: The conclusion of the project will culminate in a final report that gathers project findings, uncovered information, and recommendations for digitization tools and environments.

- Publish final report on project findings.

The work plan assembles a number of projects into a tool set for digital repositories. Tools and documentation will be released throughout the project, culminating in the release of a final toolset. These tools will proceed on independent timelines and function independently so that the community may benefit from and provide feedback to this work throughout the process. Please see appendix for more technical descriptions of *Quality Control for Video Preservation* activities.

F. Staff

Bay Area Video Coalition

Preservation Program Director: Moriah Ulinskas. (20% FTE dedicated to this project for two years) Ulinskas provides leadership, vision, and direct management of BAVC's Preservation program, one of the only nonprofit audio/ video preservation programs in the United States. Coming from a background in video art, Moriah has taught courses at the California College of the Arts, Otis College of Art and Design, and the San Francisco Art Institute. From 2007- 2011 she served as the Principal Investigator for BAVC's two National Science Foundation funded projects. She will be responsible for project management of all project activities. Ulinskas currently serves as the Chair of the Diversity Committee for AMIA.

Preservation Specialist: Lauren Sorensen. (60% FTE dedicated to this project for two years) Lauren is a 2007 graduate of NYU's Moving Image Archiving and Preservation Program. Until her recent move to BAVC, she has worked for 3 years as Assistant Director at Canyon Cinema. She has also worked as a Research Assistant for the NDIIPP project, as well as advising on a number of independent film and video collections. Sorensen oversees BAVC's preservation and digitization activities for the SMN, as well documentation and technical cataloging. Sorensen serves as the Co-Chair for AMIA's Independent Media Committee.

Information Technology Consultant: Mark Hellar. (Five hours per week dedicated to this project for two years) Mark is a consultant on technology initiatives at BAVC and number of cultural institutions throughout the Bay Area and beyond, and the owner of Hellar Studios LLC. Before opening his own studio in 2009, Mark has worked as a systems architect at the Tides Foundation, academic technology manager at the San Francisco Art Institute, and as a digital-media specialist at the BAVC. Hellar specializes in creative yet practical digital-media and web-based solutions to the technologically demanding problems faced by multimedia artists and digital-culture makers whose work requires innovative infrastructures for archiving, documentation, and exhibition.

Independent Consultants

Senior Consultant: Dave Rice. (28 days dedicated to this project over two years) Rice will coordinate, design, and manage standards, workflows, and tools to fulfill project deliverables according to archival best practices, gathered feedback, and identified technological gaps. In working with digital samples of digitization errors or irregularities, Rice will develop methodologies to automate the detection, diagnosis and qualitative reporting of digitization errors in order to ensure greater validation and responsiveness within the digitization process. Rice will oversee development of metadata standards, such as for storing frame-level checksums or quality control data. Working in collaboration with other partners, Rice will author ingest

processing scripts, metadata transformation documents, and related code. He will contribute to the project's overall documentation, reporting, and presentation. Rice serves as Co-Chair for AMIA's Open Source Committee.

Technical Consultant: Devon Landes. (23 days dedicated to this project over two years) Landes is a consultant and archivist working with commercial and government organizations in transitions to digital asset management solutions. Landes' experience with technical documentation, archival training, and quality control will support the issue tracking, documentation, and outreach objectives of the project. Landes will coordinate issue tracking and project status through development and serve as a project manager of the programming team. Landes will also design and lead the authoring of technical documentation in support of the tools, standards, and workflows enabled through the project.

Programmer and Technical Consultant: Mark Heath, SiliconTrip. (27 days dedicated to this project over two years) Heath is a computer scientist with a passion for two-dimension imaging and video and specialization in digital signal processing. Heath is a frequent collaborator with open source media projects, writing and publishing code to extend functionality of the mjpegtools and ffmpeg projects. Heath has worked within large scale video digitization and transcoding projects for on demand media companies and specializes in quality control and analysis of digital video processing.

Programmer and Technical Consultant: Jerome Martinez, MediaArea SARL. (40 days dedicated to this project over two years) Martinez is the founder and managing director of MediaArea.net, a company dedicated to creating digital audiovisual analysis tools and to open source development with extensive knowledge of digital audiovisual file and technical metadata structures. Martinez developed the open source digital media assessment tool MediaInfo, a key technical assessment tool for audiovisual data. Martinez will implement programming of tools in fulfillment of deliverables, generate software interfaces, facilitate code releases, and serve technical adviser to the ongoing project.

Dance Heritage Coalition

Executive Director: Libby Smigel. (10% FTE Year 1 and 5% Year 2) Smigel has experience in grants management dating to work in the 1970s at the NEH and has recently acquired expertise in preservation and "technology & the humanities" programs. Since 2009, she has served the DHC as Executive Director, and she has taught and published on dance history for almost twenty years. She will provide administrative and fiscal oversight for the DHC contributions to this project, will coordinate selection of tapes (diverse format and recording conditions) for the project, arrange training and pay for digitization hub Fellows (who are assigned digitization and R&D duties in the DHC satellite digitization stations), participate in dissemination, and provide reporting for the DHC outcomes and financial report.

Project Consultant: Tanisha Jones. (5% FTE dedicated to the project over two years) Since December 2009, Jones has been the DHC's project manager for the Secure Media Network, using her deep expertise in the preservation of moving images to guide the project and troubleshoot where needed. She is Director of the Archive of the Moving Image in the Dance Division of The New York Public Library for the Performing Arts, for which she selects and

assesses materials for preservation, manages rights and permissions for footage requests, prepares materials for climate-controlled storage, among other duties. Jones will facilitate and manage the project and coordinate between technology partners, the project's advisory board, and the Dance Heritage Coalition staff.

Advisory Board

The project Advisory Board is composed of preservation technology experts who will provide ongoing advice and oversight to *Quality Control Tools for Video Preservation* to ensure that the project goals are met and that national technical specifications and best practices are applied.

Tanisha Jones- Director, Archive of the Recorded Moving Image, Jerome Robbins Dance Division at The New York Public Library for the Performing Arts

Libby Smigel- Executive Director, Dance Heritage Coalition

Christopher A. Miller- Curator, e-kiNETx and Cross-Cultural Dance Resources at the Herberger Institute for Design and the Arts, Arizona State University and Dance Heritage Coalition Board of Directors

Hannah Frost- Manager, Digital Library Systems and Services at Stanford University Libraries

Greg Wilsbacher - Curator, Newsfilm Collections, Moving Image Research Collections, University of South Carolina

George Blood- President, George Blood, L.P.

Skip Elsheimer- Founder, A/V Geeks

G. Sustainability of project deliverables and datasets

The selection of our project team is intentionally focused on archivists and technologists that foster open source software development for archival objectives. Research and lessons learned from our past and current projects have incorporated open source technology and thus furthered their development. Since using open source technology without a sustainability strategy can quickly lead to the deterioration of code, support and functionality, we are looking closely at major open source projects that provide the support necessary for this project. BAVC has collaborated with developers of MediaInfo and ffmpeg to document feature requests and bugs, and to propose updates that enable these applications to be more suitable for preservation work. From these relationships we hope to share as much code from our project as feasible with relevant existing open source projects to extend support.

BAVC will make all original and incorporated code within the project available under an open license. During the project development, BAVC will make early drafts, beta versions and ongoing releases publicly available through github.com, a leading software repository.

BAVC has already benefitted from the prototyping and research on which this proposal is based, and considers the expansion and dissemination of this work of urgent importance to the field. We expect that as this project unfolds we'll have a strong user base in archives and collections.. The nine participants listed earlier who have volunteered to pilot and use the *QC Tools* at their respective organizations will ensure that the data and open source toolset continue to grow and remain relevant. Because operating systems and hardware shift as new standards in practice develop in the computing world, we want to ensure that our software tools keep up with these

shifts. These diverse testing grounds will enable BAVC to respond with accuracy to changes in protocols and usage that inevitably happen in the preservation community.

As an ongoing user of this toolset, BAVC will maintain and where necessary initiate further development of the toolset for the archival community. While BAVC will maintain the project we will also seek as the project's ongoing administrator so that submitted changes and patches from users or contributors via the GitHub repository may be considered and accepted where beneficial to the users. Potential sources for ongoing phases include the (b) (4), (b) (4), the (b) (4) – all funders that have already supported BAVC's work and are invested in video preservation.

H. Dissemination

Dissemination of the new quality control tools will be a high priority. Drafts and final versions of the proposed new tools, documentation, and guidelines will be released on an ongoing basis for review and application under an open licensing model, completely free and open for adoption or customization. Included in the budget are funds to train no fewer than five teams of peers and partners at BAVC's training facility in San Francisco on the new software tools. In addition to one-on-one guidance the trainings will be supported by and accessible to all through online training and instructions.

The project will publish documentation, training materials and software on an ongoing basis (both in draft and final form), which will be available to professional communities and the public. Users will be encouraged to see and comment on the project as it progresses and enter suggestions for its further development. This will be accomplished through a dedicated project site at a subdomain of bavc.org, which will contain references to all versions of software, documentation, and other project products with added capabilities for communication and feedback. The project managers believe in openness and collaboration throughout the development process.

Original software and source code developed within this project will be maintained at GitHub. The project will continue to use GitHub as the public repository to the Secure Media Network's openly licensed additions to the WNET PBCore Repository. Other open source software that may be further developed as part of the project, such as mediainfo and dvanalyzer, will continue to be available at their existing sourceforge.net sites and be referenced from the project site.

BAVC and partners will present the findings of this research and the development of these tools at meetings of professional organizations, such as: the American Library Association (ALA), Association of Moving Image Archivists (AMIA), International Association of Sound and Audiovisual Archivists (IASA), Independent Media Arts Preservation (IMAP), Society of American Archivists (SAA), and Theatre Library Association (TLA). At the end of the project, BAVC and its partners will develop a complete account of the project and its findings, and will make the paper available on the NEH, BAVC and DHC websites.

Notably, BAVC and Dave Rice have already been invited to present project findings on the SMN at the UNESCO International Conference, Memory of the World in the Digital Age in September 2012 which will set the stage for future presentations on *Quality Control Tools for Video*

Preservation. The sharing of findings with all the above associations, with their respective reaches throughout the library/archive, performance studies, and moving images fields, will ensure the timely distribution of the project's discoveries and tools.

In sum, *Quality Control Tools for Video Preservation* will disseminate its findings by: 1) Informal sharing of the project's products throughout the process; 2) Targeted training in tool application for no up to ten of BAVC's peer organizations. 3) Presentations at professional conferences. 4) A formal white paper to be posted on BAVC and DHC and other websites, and available to the NEH.