

Sustaining Public Media Archives

Summary of Sustainability Discussion Hosted by CLIR and WGBH

November 2017

After more than a half century of publicly supported radio and television, it is time for a nationally coordinated effort to ensure that the public's investments in news, culture, science, and educational programming are protected. Due to the deterioration of media and increasing obsolescence of older recording technologies, the American public is facing the imminent loss of much of this legacy. By involving producers, funders, archivists, technical experts, scholars, and the public, we can act now to save the parts of this legacy that are most important the future.

In November 2017, the Council on Library and Information Resources (CLIR) and WGBH convened experts and stakeholders to articulate the arguments for archiving and preserving the public media legacy, assess the challenges to doing so, and explore ways to address those challenges.¹ Participants outlined the current state of archiving efforts, discussed diverse use cases for a public media archive, and explored potential models of public-private partnerships that could address the goals of preserving legacy media, providing access, and preserving born-digital content. Following are some of the ideas discussed during the meeting.

WHY A DIGITAL ARCHIVE OF PUBLIC MEDIA?

National coordination to preserve and sustain access to public media is vital because:

- These media assets are unique sources of historical information, offering local perspectives that are otherwise submerged in narrative accounts that arise from centers of power;
- They enable media-rich teaching and scholarship;
- They serve as a historical record to be read and understood in the context of other historical materials;
- They are, in aggregate, a large data set suitable for computational analysis, enabling new discoveries and insights into historical, linguistic, and cultural trends;
- The breadth, depth, and variety of public media content supports inquiry in humanistic, social scientific, and scientific disciplines; and
- They need to be accessible to the broad public, and not just locally, or in limited collections, but as a whole collection from across the country.

¹ The meeting agenda and a bibliography are available at <https://clir.wordpress.org/initiatives-partnerships/sustaining-public-media>.

WHAT ARE THE CHALLENGES?

To sustain a public media archive, a coordinated effort must overcome three “grand challenges”:

1. Digitization: A huge amount of at-risk public media resides in hundreds, if not thousands, of organizations throughout the United States. There is an imminent threat of permanent loss due to deterioration of materials, often exacerbated by poor storage conditions. Determining the most significant at-risk public media that should be digitized, raising significant funds to undertake this work, and coordinating this work at a national level are all necessary steps toward saving these materials. The scarcity of equipment suitable for transferring content from obsolete technologies and the dwindling numbers of people with the expertise to use that equipment are additional factors that make the migration of legacy public media an urgent priority. While technological advances may reduce costs for other preservation activities, the costs of digitizing obsolete audio and audiovisual formats will increase over time, making immediate action to protect past investments in public media financially prudent.

2. Access: Digitization is a key step in saving legacy media, but by itself does not render the material accessible to the public or useful to scholars. Additional work is required to make content accessible and useful to communities of interest having varied needs, interests, and skills. Digital audio and audiovisual files must be susceptible to the kinds of tools for discovery and analysis that users wish to employ. Information professionals, user communities, and technologists must work together to identify the minimal necessary and viable metadata requiring human labor, and also further develop automated indexing methods that will enable communities to find and use legacy content effectively. Access efforts must proceed in conversation with public media creators so that intellectual property interests are considered when content is stored for archival access at a later date.

3. Preservation: A digital archive of public media must preserve and make accessible a massive and ever-increasing wealth of both newly digitized and “born-digital” content. Through automated indexing and metadata capture, advances in technology can significantly reduce the costs of ingesting new content. To realize these savings, it is critical that public media archivists collaborate extensively with public media creators to integrate archival practices into production workflows. The declining costs of storage may allow more to be stored digitally at current costs, but digital collections are growing at an exponential rate. However, long-term preservation also requires regular migration to updated formats, updated software systems, and even updated hardware. The archive must establish a structure that is future-compatible, enduring, and self-sustaining.

WHAT ARE THE OPPORTUNITIES?

Despite these daunting challenges, now is an ideal time to act to preserve and sustain a public media archive.

Opportunities Around Technology

Opportunities arise from falling costs and the potential for developing revenue.

Falling Costs

- Advances in storage technology and the creation of new storage capacity, both driven by commercial interests, are reducing the cost of storage and will continue to do so, making it possible to accommodate the ever-increasing volume of public media content over time, assuming ongoing access to sufficient funding.
- Advances in such areas as image and audio processing, speech to text, facial recognition, and automated tagging make it possible to automate much of the processing required to render digitized archival content manageable and useful. Making large quantities of digitized public media accessible is more affordable now than it has been in recent years and should become even more affordable as open source tools that facilitate this work become more refined.

Revenue

- Although access to archived public media recordings should remain a free public service, the development of open source systems and processes to sustain an archive could enable opportunities for applying those systems and processes to commercial purposes, generating revenue for the archive.
- The archive can look to the models of those who have successfully built businesses on the development of open source tools, the uses of which were then commercialized.

Opportunities Around Partnerships

Sustaining a public media archive creates opportunities for the formation of new interdependencies.

- A public media archive will require contributions and ongoing support from researchers and stakeholders from the educational community in addition to life-long learners, professionals such as journalists and writers, and consumers such as genealogists.
- The archive will better enable local media stations to identify, engage in, and develop ongoing relationships with their surrounding communities as well as enable more effective collaboration with other stations across the country.
- The archive can enrich reporting and inform journalists, policy makers, and the public on current topics of concern, thus helping advance conversations on these topics.

Opportunities Around Sustaining Access

A coherent, comprehensive public media archive built over time will advance and accelerate the creation of new knowledge.

- Successful research and learning requires coping with an ever-increasing volume, rate of production, and variety of content—the massive amounts of unmediated and uncurated data that are being generated daily. Public media is an important driver within this stream of information.
- A functional public media archive can serve as a testing ground for researchers to develop tools with which to mine the data of the live web, while wider trends and developments in web archiving can be harnessed for the benefit of capturing and sustaining public media.
- A growing collection of publicly accessible and reusable metadata describing public media content will itself be a valuable resource for research.

Opportunities Around Legacy Content

While the need for resources to preserve legacy content is enormous, there are diverse constituencies interested in media preservation that should be receptive to a nationally coordinated campaign.

- Digitizing legacy public media remains expensive and, like public media production, requires ongoing direct investment.
- Owners of legacy public media content should seek to combine digitization efforts into larger-scale projects to drive down the costs of digitization.
- In recognition of the imminent risk of lost content and the need to set priorities for preservation, owners and holders of legacy public media should not face criticism if they choose to let go of low-priority collections, so long as they relinquish them to others willing to “adopt” and preserve them.
- The urgency of the situation requires an evangelical, orchestrated campaign to raise awareness and concern, and to identify possible adopters of orphan or released content.

Opportunities Around Born-Digital Content

Effective strategies for ingesting new public media content into an archive will benefit the archive, its users, and producers alike.

- Born-digital content is fragile and susceptible to loss without thoughtful care, but it is infinitely “copy-able.”
- When archivists and media makers establish partnerships of trust, they can plan for archiving at the point of creation, building “archive-readiness” into production workflows. “Archive-readiness” could be supported by building appropriate tools and platforms.
- Funders and national organizations supporting public media can encourage or even require archiving at the point of creation in situations where there is sufficient leverage to do so.

- “Archive-readiness” can also facilitate upstream use in a time when content producers are ever more aware of the ongoing value of their products. Better archiving will result in more effective asset management.

NEXT STEPS

Acknowledging and accepting these challenges and opportunities, CLIR and WGBH invite the community to join us in these actions:

Evangelize: We need leaders who can engage effectively with communities of producers (media creators and funders), consumers (educators and the public), and collectors (libraries, archives, and museums) about the three “grand challenges” and the opportunities they present for immediate action. We need communicators who can capture and keep the public’s attention on the problem of lost legacy and the potential steps individuals, groups, and organizations can take toward addressing it.

Prioritize: We need easy ways to identify at-risk content and a set of tiered options for preserving it. Public media archivists, producers, and stations should know what it takes to minimally capture and preserve content that can be accessed later; what it takes to capture and make it minimally discoverable; and what it takes to capture, make discoverable, and facilitate reuse of such content. Those who have content they cannot afford to preserve need a way to find partners who can “foster” or “adopt” the content for preservation reformatting.

Analyze: We need more information on producers’ working practices in order to identify easy ways to implement archiving practices at or near the point of creation. We need more information on strategies for lowering barriers to access to legacy content so that more can be made available to the public, sooner.

Collaborate: We need new partnerships focused on the three “grand challenges” facing public media archiving—reformatting legacy content, implementing long-term preservation strategies for both legacy and born-digital content, and surmounting barriers to access.

For more information, contact Karen Cariani, karen_cariani@wgbh.org



Council on Library and
Information Resources