

Report from Preserving Digital Sound and Vision: A briefing, London 15th April 2011

Delaina Sepko, University of Glasgow

1. Introduction

This document reports a DPC members' briefing day held in London on 15th April with the title 'Preserving Digital Sound and Vision'. The purpose of this report is to assist DPC in measuring the success of its events programme, help shape future work in the field for the DPC and provide a commentary and narrative for those DPC members unable to attend. It includes a narrative report drafted by Delaina Sepko who has also compiled evaluation from participants.

2. Preserving Digital Sound and Vision: A Briefing

Emerging tools and services for digital preservation are typically built around the need to preserve texts, documents, images and data sets. Audio and video – broadly defined as time-based-media - have received less attention within the library and archive communities, partly because they have historically been seen as distinct, partly because they present new technical challenges, and partly because they have hitherto represented only a small proportion of the collections which memory institutions and research archives collect. However, the simplicity with which digital video and audio can be captured and the ease and popularity of online distribution means that they are now ubiquitous, creating new concern for long term access. As more and more of our cultural and scientific legacy is being created in digital audio-visual formats, so those managing long term access to data need to understand the challenges and opportunities which these formats bring. New skills and new techniques will be required to ensure our digital audio and video memory is accessible tomorrow.

This DPC briefing day provided a forum for members to review and debate the latest development in the preservation of digital sound and vision. Based on commentary and case studies from leaders in the field, participants were presented with emerging tools and technologies and were encouraged to propose and debate new directions for research. The day included a discussion of key themes including.

- Preservation of digital audio formats and collections
- Preservation of digital video formats and collections
- Creation, documentation and retrieval of digital sound and video
- Emerging tools and policies for preservation

3. Narrative Report

- **Nicky Whitsed, Open University Library Services**

In the past, AV collections in libraries have tended to be neglected. Traditionally, this material has been catalogued using library standards, which seldom make provisions for appropriate AV metadata. Metadata relating particularly to IPR and ownership is key when deciding how to make AV collections available to users. In order to make digital content available, both technical and IPR challenges should be addressed. A collection's value, whether for business or heritage purposes, is often compared to the price of preservation or re-creation. Finding innovative ways to re-use AV material can produce a strong argument in favour of preservation. Indeed, ways to effectively re-use AV material need to be considered during creation and these methods need to align with institutional objectives as well as take into consideration inter-institutional content strategies and storage solutions. Who will decide what to save? A digital preservation policy will not only be

important for deciding what and why material is saved but will also provide consistency across an institution and over time. A successful policy will have the support of both management and users and will identify risk in 'do nothing', 'do something' and 'do everything' preservation scenarios.

- **Richard Ranft, British Library Sound Archive**

Decay and loss of analogue content can be delayed but cannot be prevented for long, so digitization is a key element in the preservation of any sound collection: and this creates a digital preservation challenge. Many of the technical issues relating to digital preservation of sound collections have been solved. There are a few established best practice guidelines and standards such as Sound Directions that offer advice for digital preservation. This all goes to make preservation through migration a necessity. Digitisation is time consuming and can be labour intensive so it should be done once to the highest standards. Digitisation is only part of the preservation process; playback machines, their replacement parts and the skills to operate and fix them need to be researched. Setting up digitization from specialist analogue recording and play back equipment can be much more time consuming than simply playing and recording the sound. Effective preservation should produce three files for different purposes: un-manipulated master files, cleaned up playback or working files and small, easily accessed user files. While navigating all stages of digital preservation, you should expect issues relating to time, workflow efficiency and legal restrictions.

- **Stephen Gray, JISC Digital Media**

There is too little research that identifies the specific users of digital content and their needs. Evidence shows that an increasing number of researchers are engaging with AV material in their research and JISC Digital Media have attempted to evaluate this particular group of users' needs through interviews, surveys and analysing existing user statistics. This study revealed that researchers share four needs: citation, acquisition, analysis and publishing. Consistently citing AV material, especially internet-hosted material, can be difficult and there is no guarantee that media will be there tomorrow. Currently, reproducing content for educational or archival purposes can create legal difficulties and IPR issues still present challenges. Any media produced during research should be fit for purpose and its significant properties must be identified. Furthermore, the analysis of any digital content should consider ethical and legislative issues. A small but growing amount of research is published in a cross-media format, which is now recognised by the REF. Despite an increasing engagement with AV media in research, digital preservation has not trickled down to the average researcher. For more information on this research project or advice to creators, visit www.jiscdigitalmedia.ac.uk.

- **John Zubrycki, BBC**

The standards and formats for broadcast and audio content are varied and present a number of challenges for preservation. One of these challenges is integrating digital file based operations into digital production workflows. The BBC has developed a tapeless production workflow, which has been designed to compliment its Digital Media Initiative. The DMI aims to digitally convert all of the BBC's AV content and improve the ways digital material are made, shared and used. In this workflow, for example, creators can easily ingest, edit and add metadata to digital AV content. Traditionally, archives preserve content in whatever form it was given to them, but digital files separate content from its media. Therefore, compression can be applied to content regardless of its media. In broadcast situations, compression should be applied cautiously so that content is visually lossless or it should be avoided. Compressed AV content is more vulnerable to errors, but it is common for any broadcast content to have three layers of compression applied during creation, storage and transmission. The variety of encoders used to compress content raise concerns about their sustainability. To help relieve these concerns, all archives should adopt the same or same family of encoders. Compression is unavoidable because it helps manage the vast amounts of data being stored, but any storage and file management systems should be scalable to accommodate

future holdings. Data integrity checks should work at content level and by means other than listening or watching. PrestoPRIME, a research project with the goal to develop a preservation framework for AV content, addresses these issues. For more information, visit www.prestoprime.org.

- **Richard Wright, BBC**

There are dozens of preservation tools for AV material. Unlike analogue content, the ‘do nothing’ approach to digital content does not work, although elements of an analogue preservation workflow will be helpful. In fact, there will be several stages in digital preservation. Digitisation is only the first stage of preservation. The second stage addresses the workflow, processes and storage requirements. The third stage addresses the mechanics of preservation – migration, OAIS, obsolescence etc. Progressing through these stages, we must ask how or if text-based digital preservation strategies apply to AV content. AV material is time-based, but it is still subject to traditional concerns about handling, environment and its location within a trusted repository. It also has similar access requirements – granularity, navigation, citation and annotation. The Presto Centre, which includes PrestoPRIME, aims to answer the ‘what if’ questions like those concerning storage systems, storage configuration and file collections raised during digital preservation. For more information, visit www.prestocentre.eu.

- **Sue Allcock and James Alexander, Open University**

The Access to Video Assets (AVA) project based at the Open University has developed a repository prototype for managing the digital preservation of video. Its challenges have been bringing together different collections that have different functions, accommodating various formats and the inconsistencies of referencing videos across collections. AVA has addressed these challenges by identifying and locating at risk tapes, their metadata and paperwork and user needs. AVA has consulted users who want videos preserved and available for re-use. As is the case in many repositories, OU’s videos are only available to staff because of difficulties with IPR. In order to meet user research needs, AVA has identified a need to locate particular points in a video; simply directing users to a two hour video is no longer suitable. Meeting user needs, the OU developed AVAFedora, which uses video subtitles to improve finding and direct users to specific points in the narrative. Embedding preservation into workflows and encouraging access to content have been important arguments to the project’s case for preservation, which has always been driven by user needs.

- **Simon Dixon, Queen Mary University of London**

Currently, users can find a relevant document in a collection through bibliographic metadata, free text descriptions, queries or searches and recommendations. How can users search and find information for time-based content like music? Options like placing tags at significant places in a song or providing fast forward/rewind capabilities will be determined by how users are engaging with music. Audio retrieval systems have introduced audio semantics, which draw out ‘meaning’ from audio signals. This meaning can range from descriptions like a song’s key signature, tempo or tuning system. Retrieval systems can also draw upon annotation of significant features of a song such as instrument entrance, song structure, beat tracking and automatic transcription. Ontology-driven Music Retrieval & Annotation Sharing service (OMRAS2) is a framework that uses standardised descriptions for online music metadata to improve user searches in time-based media like music. The Sonic Visualiser, OMRAS2’s open source desktop application, allows the viewing and editing of metadata and supports extra plug-ins for features like beat tracking. For more information, visit www.omras2.org.

Discussion points for DPC

Storage

- How confident are we about destroying media after it has been digitised?

Sometimes our backs are against the wall; technology obsolescence or lack of skills or knowledge will decide this answer for us. Yet there is a resistance to throwing out media because it may have valuable analogue assets associated with it. If destroying the media is an option, then be sure you have lots of digital copies. It need not be an all or nothing situation though; remote storage is an option if material is not regularly used.

- Is digital storage getting cheaper at the same rate as the demand for it?

There seems to be a stepped demand that reflects shifts in technology i.e. the amount of space needed to store standard verses HD video. Archives do select content so they probably will not save everything they create or acquire. This point was countered by the question 'if the space is available, why not save it all?' Ultimately, we will need to ask how the cost of appraisal compares to the price of storage. Determining storage cost is not the answer to the question; we must also consider the cost of access and retrieval.

Rights Management

- Are there any other regulatory or rights issues that need addressing?

Do companies like Google even care about this question or can they afford to take the risk? These companies may host or post content to attract users, but the 'take down' policies may be more interesting to look at. The type of media may make a difference when asking this question; the Gowers Review still does not address access rights such as making preservation copies of recordings. The cost of gaining permission to use AV material can be very high. It helps to use collection agencies because it cuts down on the number of individuals that need to be contacted. There was a general consensus that there need to be educational and archival exceptions made in IPR legislation.

The next step

- What would be helpful in moving forward with digital preservation of sound and vision?

The key points raised were how to retain AV contextual information, how smaller institutions can pool resources and how institutions can potentially receive support from BBC R&D's Ingex. It was suggested that maybe the DPC could act as a liaison for this support as well as promoting best practice for digital preservation.

4. Evaluation by Participants

Basic Facts

47 people attended the session; 30 completed an evaluation form.

Type of organisation:

Library	Archive	Museum	Other (specify)	No response
8	11	3	16 Educational institution, MAM development, Police, Broadcast R&D, Consultancy	0

Role:

Librarian	Archivist	Conservator	Other	No response

2	10	5	18	0
			MSc student, Information Officer, Academic Researcher, Artist, Owner, Digital Preservation Architect, IT, Senior Strategist, Trainee, Engineer, Advisor, Project Manager, Director, Research and Development, Digital Preservation Developer, Technician, Systems Manager, Developer	

What were your reasons for attending this event?

- Audio preservation is one of my research interests.
- Expand knowledge; hear latest thinking, case studies.
- At Tate, we collect and preserve artworks that are born digital audio and video.
- Invited speaker
- Practical knowledge; to catch up with progress of tools and services I know about from previous projects; to inform my thinking on current projects.
- Network, learn and contribute
- Learn more about AV side of digital preservation.
- Information gathering for local audio/visual storage and delivery project
- To learn about ... now and the future.
- Meet peers, hear other viewpoints, survey the landscape
- Awareness and professional networking
- Setting up a digital preservation facility and audio digitisation within Tate Archive.
- I am working on a project to preserve digital video files.
- Help to spread technology solutions to a wider audience.
- More inter-sector discussion of such a specific topic is great.
- Raise awareness of digital preservation tools
- Invited speaker
- Interest in AV preservation. Background for Tech Watch Report.
- Gather information about preservation and digitisation of AV material for a new project.
- We will be getting more and more AV material. I'd like to know how best to manage it.
- Looking to undertake a major AV digitisation project within the next year.
- To exchange ideas and encourage working together.
- Interest in AV preservation techniques and trends, new tools, migration for preservation formats.
- To find out more about what is going on in other organisations regarding preserving and providing access to AV files.
- To keep up to date with developments in this field.
- To understand issues of digital preservation of AV material.
- Research regarding Digital Asset Management and time-based media conservation.
- Interest in preserving audio.
- Introduction to AV preservation.
- Exchange ideas, discover best practice, disseminate project outcomes.

On a scale from 1-5 how would you rate today's event (actual numbers):

	Not satisfied				Very satisfied	No answer
	1	2	3	4	5	
Relevance to you	0	0	0	8	10	12
Presenter(s)	0	0	0	6	12	12
Level of information	0	0	0	11	7	12
Venue and facilities	0	0	2	7	9	12
Value for money	0	0	0	2	14	14
Overall satisfaction	0	0	0	6	12	12

On a scale from 1-5 how would you rate today's event (% of answers received):

	Not satisfied				Very satisfied	No answer
	1	2	3	4	5	
Relevance to you	0	0	0	44.4	55.6	0
Presenter(s)	0	0	0	33.3	66.7	0
Level of information	0	0	0	61.1	38.9	0
Venue and facilities	0	0	11.1	38.9	50	0
Value for money	0	0	0	12.5	87.5	0
Overall satisfaction	0	0	0	33.3	66.7	0

Which sessions did you find most useful? (multiple choices possible)

Defining the 'Value Proposition' for Audio-Visual Preservation: The Nature of the Problem, Non Scantlebury, Open University	4
Collecting, conserving and managing sound archives Richard Ranft, British Library Sound Archive	5
Creators and consumers: digital sound and vision as research outputs, Stephen Gray, JISC Digital Media	6
Challenges and solutions in broadcast archives, John Zubrycki, BBC	13

Presto Centre and emerging tools in preservation of audio-visual collections, Richard Wright, BBC	12
Access to Video Assets Project, The Open University: Outcomes and Lessons Learnt, Sue Allcock, James Alexander, Open University	9
Emerging tools for digital sound Mark Plumbley and Simon Dixon, Queen Mary University of London	3
Discussion: what is to be done, why and by whom	6

What did you learn? What will you do (differently) as a result of attending this event?

- How organisations are approaching digital preservation.
- Importance of selection, cost, storage. Discussion of use, access and value.
- Yes! Learn more about checksums and fixity. Try out the tools mentioned by Simon Dixon and Richard Wright.
- I'm new to preservation and I learned a lot.
- Awareness of current research in the field will inform my current and future work.
- The wheel is being reinvented!
- AV digital preservation field has already solved many of my non AV problems.
- Presentations at this session will inform requirements for gathering process of our video storage and delivery project. Insight of various areas that need to be considered, into existing tools.
- Where there are sources for advice and guidance, useful case studies, contacts.
- Some tech info, but more about other approaches, who is doing what, what consensus views might look like
- 'Industrialisation' of digital preservation - simplified, standardised and cost effective.
- Carry out more detailed assessment of access needs and explore potential of automated metadata creation.
- I think I have more ideas about how to approach my work and more ideas of places to look for help.
- Work more closely with JISC and OU on standards and best practice
- Lots of 'reading' to do. Will stay in touch with DPC and Presto web resources.
- Assisted greatly with my understanding of digital preservation and associated problems various colleagues have to contend with.
- Good summary of state of AV preservation. Will help in final project report. Good examples for benchmarking.
- Notes for Tech Watch Report
- Far too much to note! General principles, good practice, digitisation and software.
- Will have more of an idea of how to deal with AV when we take it on.
- Searching long video and audio clips - better ways to do this.
- Others have similar problems.
- Good to hear about standardisation processes in archiving. Presenting disparate formats. Really interested in tag means of navigation and jumping around into marked points of interest in AV files.
- I'm very interested in ways to make videos and audio files searchable and accessible by adding metadata and labels.
- Preservation of file formats for digital AV material. New tools becoming available.
- I'm not the only one with preservation problems.
- A mass of new information.
- Audio tools of great interest.

Was there anything else you would have liked us to have included?

- More chat about capturing and preserving social context of AV content.
- A bit more on metadata, but this only a minor comment.
- Can't think of anything!
- The challenge of digital platforms including social media and networks. How do you capture and preserve AV in context? Eg. The website or means of publication and dissemination of AV.
- I missed some presentations so I would have liked to see these.
- Ideas of software used to rip DVDs. Anyone who has recently had to deal with managing their digital video resources.
- Service providers: the commercial answer to digitising small collections.
- Practical steps for smaller collections?
- No
- More use cases, if any?
- More veggie sarnies!
- More about AV (non-BBC related).
- Video recording of the talks for later viewing.
- More!

What did we do well?

- Everything. You're great hosts and clearly make a wide range of people feel welcome.
- Everything. Thank you.
- Everything!
- range of presentations and summary discussion
- Coverage of a wide area of DP requirements related to AV.

- Good range of subject areas but limited presentations from different sectors. No speakers from regional AV archives or local authorities? Community archives? Eg. Can DPC support smaller collecting institutions and bodies?
- The whole day seemed to run smoothly and efficiently.
- Provide a range of sessions addressing digital preservation as an end-to-end process.
- Kept to time well. A range of topics.
- Very good mix of information and discussion.
- Mix of speakers and provide an overview.
- Excellent networking opportunities.
- Organisation, debate, range of speakers
- Good discussion session.
- Varied in level of discussion - scary but useful! Awareness. Excellent demos from presenters.
- Venue and range of speakers.
- Hosting discussions
- Everything
- Discussion session at end was very useful.
- BL is a good venue.
- The debate.
- Great pace of presentations and a lively debate.
- Event well-paced, great location and good discussion.

Where did you hear about this event?

Twitter	1
Digital preservation list	10
DPC discussion	3
By Invitation	10
DPC Website	2

Other: 3 – Seminar; Recommendation; Unspecified

Thinking about everything you've heard today, what do you think will be the key challenges in preserving digital sound and vision? How and by whom might it best be addressed?

- IPR issues. Cohesive and coordinated discussion with legislative bodies.
- Getting together a working group to help the smaller organisations understand when to out source, when not etc.
- Working in collaboration so that we bring about change that serves the maximum amount of stakeholders and also so that we don't reinvent the wheel.
- Budgets. Explore business models or cases.
- storage and legal
- Cost (storage, rights clearance); consistency of AV file formats, codecs etc across institution; consistency of metadata.
- The exponential growth in digital AV and the challenge of search and navigation. Standards to allow federated access and search across separate collections.
- Digitisation during window of opportunity (degradation/obsolescence). Persuading wider population to support funding of large-scale digitisation projects. Preservation of the bits. Preservation of the meaning of the bits.
- Rights issues and the ability to capture the full spectrum of born digital material.
- Technical knowledge. Technical obsolescence. Money and resources. Sharing knowledge and procedures between institutions.
- Budget - so work on access and examples of institutional value need to be collected and promoted.
- Scaling solutions without 'ditching' best practice.
- Given me a better understanding of who should be involved in preserving digital sound and vision.
- Tapeless/online/web preservation.
- Rights - DPC and associated bodies.
- How best to describe how to meet researchers' needs.
- Changing carriers. Formats need to be changed into different formats.
- How archives can work together. All of us, with the help of the DPC, should address it.
- High level of storage and enabling access on demand to any individual file from any location. How does anyone know what is available to see and hear? Are there more intuitive ways of searching - amazon approach?
- Search and retrieval of resources. Citing resources.
- Codecs, storage choices and volumes.
- Analogue digitisation
- Rights management.
- Budget and difficulty curating born-digital content.

Any further comments

- Well done. I really enjoyed myself.
- Very good. Many thanks.

- Thank you. It was really well organised and all the talks were excellent.
- Thanks!
- Let's get together with the Tate, BBC, JISC, DPC and ourselves to review options for Ingex.
- The presentations were more focused on search and retrieval eg the use of digital AV rather than the preservation. Does that reflect a changing need in the AV community? Almost all speakers needed 30 minutes to present.
- Greatly enjoyed opportunity to meet peers and discuss digital issues at a strategic level. Wish I'd been around for more of it.
- Very good event.
- Great day. Thanks.
- Thanks for a very interesting and stimulating day.

5. About this document

Version 1	Document collated from submitted reports	18/04/2011	DS / WK
Version 2	Comments and corrections from participants
Version 3	<i>Distributed</i>	21/04/2011	..