

# Preservation of Broadcast Archives – a BBC Perspective

by Richard Wright, Senior Research Engineer  
BBC Research & Development, UK

This paper covers: 1) the basic preservation problems of audio-visual archives are reviewed, including data from the PrestoSpace and TAPE surveys; 2) an approach to preservation planning, stressing that preservation work needs to be informed by a general strategy for a collection or institution; 3) the work needed to build a business case in order to get funding; 4) the general principles of audiovisual preservation; 5) a detailed roadmap showing the preferred strategy and choices for migration from old formats, and finally 6) the new problems of digital preservation.

## Broadcast content: who cares?

The tomb of John Keats in Rome has the inscription: "Here lies one whose name was writ in water". Broadcasting has a similar concern – our output, the whole result of all our effort – is writ 'on the ether'. There is nothing left of broadcasting unless it is recorded, collected and preserved; an operation that would be sterile unless there are also mechanisms for access to the preserved content.



Figure 1. Keats' Tomb. <http://www.flickr.com/photos/sgrunt/2077401671/> Creative Commons licence.

The problems of broadcast preservation begin with collection. Most European countries have a tradition of public service broadcasting, and part of that tradition is the existence of broadcast archives – either national institutions as with the French Institut National de l'Audiovisuel<sup>1</sup> and the Netherlands Institute for Sound and Vision<sup>2</sup> – or archives of individual broadcasters (for example RAI, ORF, SVT, RTE, NRK, BBC – and the regional broadcasters in Germany). In the UK the situation is complicated by an overlap of concerns, shared amongst the BBC, the British Film Institute<sup>3</sup> (which is also the national archive for television content – *except* for the BBC, though the

1. [www.ina.fr](http://www.ina.fr)  
2. <http://portal.beeldengeluid.nl/>  
3. [www.bfi.org.uk](http://www.bfi.org.uk)

BFI is the public access mechanism for BBC content, and has a mechanism for recording off-air) and the British Library Sound Archive (which provides public access to BBC radio content, and does selective off-air recordings of other radio broadcasting). To add complication, access to broadcast content in higher education is governed by the Educational Recording Agency<sup>4</sup>, and TV content (BBC and other channels) can be provided to higher education from off-air recordings made by the British University Film and Video Council<sup>5</sup>.

The UK situation is awkward, but at least there is a lot of recording and collecting. In the USA, where the major broadcasters are commercial, the broadcast archives are also commercial and provided limited public access (Jeff Ubois has documented the difficulties in his paper<sup>6</sup> describing a search for TV material from the early 1980s). Odd exceptions have developed, allowing Vanderbilt University to record television news<sup>7</sup> off-air, but otherwise TV is archived by the broadcasters themselves. Public broadcasting in the USA is similarly a dispersed activity, based on individual stations or even individual programmes. There is no comprehensive preservation programme.

That's the situation as seen politically. An engineer's view asks for the technical specification: what exactly is being collected? Here there are two basic answers: *studio quality* (meaning the highest quality available) material can be collected, or recordings in *broadcast quality* can be made off-air. The latter is the easy option, especially for a national body. But the former, archiving of studio quality content, is the *only* way to capture the full technical quality of the content. As with any archiving process, it is a matter of principle to always seek to 'capture the best'. Just as one would prefer an original manuscript to a photocopy, capture off-air is a poor second-best, to engineers and to archivists.

To show the difference, here are two photographs that differ by roughly the same factor as the difference between studio and broadcast quality. If only the second is archived, the detail and correct colours of the original are lost forever.



Figure 2. High quality image (left) and low-quality, as in off-air recording (right). © RW

4. [www.era.org.uk](http://www.era.org.uk)  
5. [www.buifvc.ac.uk](http://www.buifvc.ac.uk)  
6. Ubois, Jeff, "Finding Murphy Brown: How Accessible are Historic Television Broadcasts?", *Journal of Digital Information*, Vol 7, No 2 (2006); <http://www.archival.tv/MurphyBrown-final.pdf>  
7. Vanderbilt Television News Archive: <http://tvnews.vanderbilt.edu/>

### Problems of audiovisual material

In 2004, the EC project PrestoSpace<sup>8</sup> performed a survey<sup>9</sup> covering 11 countries, with additional data from the public websites of audiovisual archives in another 9 countries. The 2004 survey covered the largest archives in each of these countries. The data was greatly extended by the survey in 2005 run by the EC project TAPE<sup>10</sup>, with responses from nearly 400 archives of all sizes.

**How much audiovisual material is out there?** The TAPE survey found about 25 million hours of film, video and audio in Europe. PrestoSpace had found 10 million hours, but there is an estimated five million hours in common between the two surveys – meaning a total of 30 million hours of audiovisual content held in formal collections. These figures support the general estimate of 50 million hours in Europe, and 200 million worldwide<sup>11</sup>.

**What condition is it in?** According to TAPE, 70% of material is seen by its curators as in acceptable, good or very good condition, and 30% is deemed deteriorating or unknown. **But half the archives do not have controlled storage conditions, half have no regular equipment maintenance, and 2/3 do not have a systematic preservation programme.**

**What is being done?** Preservation projects were planned or underway to transfer about 250 000 items per year: about 1.5% of total holdings. At that rate it would take 60 years to deal with current holdings. Much of the material will not last for 60 years; average 'format life' of videotape is 20 years or less (as little as 10), and then the format becomes obsolete. Life expectancy of the material itself varies with storage conditions, but without cold, dry storage most audiovisual materials deteriorate after 20 to 30 years. Further:

- New material keeps coming in; project Presto found that acquisitions were exceeding preservation work by a four to one ratio<sup>12</sup>.
- There is already insufficient budget and insufficient resources: the PrestoSpace survey found that archives had half the budget they needed (just for their planned 1.5% per year transfers), and the facilities providers also had half the needed capacity.

### Preservation planning

The 30 million hours of audiovisual content just referred to largely sits on shelves, and it has a limited shelf life. Format obsolescence, media deterioration and damage mean that all the video and audio will need to be digitised and moved to some form of digital storage. The situation for film is different, as the format is not obsolete in the same sense as for audio and video. Film projectors and film itself can easily last for another century, and some forms of film, in proper storage, could last 400 years<sup>13</sup>.

8. <http://www.prestospace.org/>

9. Richard Wright, *Annual Report on Preservation Issues for European Audiovisual Collections*; [http://digitalpreservation.ssl.co.uk/asset\\_arena/text/4/D22-4.pdf](http://digitalpreservation.ssl.co.uk/asset_arena/text/4/D22-4.pdf)

10. Edwin Klijn and Yola de Lusenet, *Tracking the reel world*; <http://www.tape-online.net/survey.html>

11. Schüller, D., *Making Audiovisual Contents Available*, EuroChina 2002; [www.eurochina2002.com/docs/Online%20after%2024%20May%2002/Cconference\\_Presentations/Dietrich\\_Schuller.pdf](http://www.eurochina2002.com/docs/Online%20after%2024%20May%2002/Cconference_Presentations/Dietrich_Schuller.pdf)

12. <http://presto.joanneum.ac.at/projects.asp#d2>

For all the audio and video (which was 95% of the 20 million hours in the TAPE survey), active steps will have to be taken or the material will be lost, within 20 to 30 years. This process has started, and indeed the new project PrestoPRIME<sup>14</sup> has estimated that broadcasters have digitised something like five million hours of content already. However it is an expensive process, which requires planning – first in order to know what to do and the most economical way to do it, and second in order to build up the information – the business case – that is inevitably needed in order to raise funding.

PrestoSpace has published an online guide<sup>15</sup> to audiovisual preservation, which gives more detail and more examples; this article will give a summary of the recommended planning process.

The PrestoSpace recommended process has these stages:

- **Cartography and triage:** make a map of your collections, and divide the overall task into parts, by priority.
- **Develop a collection strategy:** preservation is a strategic issue, and must fit the particular circumstances of a collection or institution. In the BBC, we are moving (gradually) from 1 million hours of content, on 100 km of shelves, to an all digital collection on various forms of digital storage. This change will be part of a major change in the ways of working across the archive, and across the whole BBC. We need to do the best we can to plan our preservation work in order to achieve an all-digital archive that best suits the future usage requirements of that archive – if we can predict them!



Figure 3. A bit of the BBC archive's 100 km of shelves – eventually all to be digitised. © BBC: photo used with permission.

- Within this collection planning a *preservation strategy* can then be developed. The preservation strategy is about much more than digitisation – because digitisation is one process and preservation has to be about all the processes needed to maintain content.
- Finally, for specific groups of material, a specific *preservation plan* will set out the actions, such as digitisation of one audiovisual format, that are identified for treating as a specific project. All the usual techniques of project management, from funding to sign-off, can then be used on individual actions in a preservation plan.

13. *Preserve then Show*, Danish Film Institute; [www.dfi.dk/NR/rdonlyres/93B04DFF-ECCE-426B-8AA9-0CFD53707B3E/0/JesperStubJohnsen.pdf](http://www.dfi.dk/NR/rdonlyres/93B04DFF-ECCE-426B-8AA9-0CFD53707B3E/0/JesperStubJohnsen.pdf)

14. [www.prestoprime.eu](http://www.prestoprime.eu)

15. Richard Wright, *Preservation Guide*; <http://wiki.prestospace.org/>

- A collection strategy should at least consider the following:
- The long-term purpose of the preservation work: where are you going with the collection?
  - Access: audiovisual archives have traditionally had very limited access, for a range of reasons (technical, legal, logistical). Now, all but the legal constraints have virtually disappeared, and there is no technical barrier to universal instant access to all your holdings.
  - Required changes to how the collection is managed: a digital collection doesn't have the same workflow and management processes as a shelf-based collection. Circulation control could disappear entirely. Acquisition of digital materials can be automated. Cataloguing may need to suit the needs of a much wider and less specialised set of users. In the BBC, we have the results of something like 2000 person-years of cataloguing effort in our TV and radio catalogue, designed to be used by librarians and professional 'film researchers'. We need a major change just to get this catalogue reshaped for public access – not to mention the effort needed to get the content itself online. Table 1 shows just one problem – our extensive use of specialist abbreviations.

<b>b/w</b>	black and white
<b>bldg</b>	building
<b>cpyrt</b>	copyright
<b>demo</b>	demonstration
<b>dep acty</b>	dependent actuality
<b>disco</b>	discussion
<b>doc</b>	documentary
<b>ex lib</b>	ex library
<b>ext</b>	exterior
<b>ftge</b>	footage
<b>fx</b>	effects
<b>gfx</b>	graphics

Table 1: The first dozen items in one BBC list of cataloguing abbreviations.

- What the preservation projects should contribute to the collection strategy: a time when whole blocks of material are coming off the shelves is a time when other actions can be most convenient and economical. Examples are:  
Installation of new shelving  
A full inventory check  
Repackaging  
Photographing or scanning the packaging  
Getting metadata off the packaging, or checking it for accuracy

Here is the BBC's preservation strategy for film:  
**Preservation Strategy: BBC 16mm film**

Type of material	Condition	Action needed	Timescale	In-house or contracted?
16m mag sound track - masters	vinegar syndrome!	digitisation to file formats; destruction of originals	2 years starting immediately	Contracted; checking in-house
16m mag sound track - duplicates	vinegar syndrome!	destruction (after respective masters are transferred and checked)	2 years starting immediately	In house
16mm Ektachrome	some colour fade	Access copies made on digibeta and DVD	Starting when budget allows: in 2 years	Preparation and checking in-house; telecine contracted out
16mm B&W film negatives	good	Maintain in appropriate storage conditions; review condition at intervals	Review plan and condition every five years	Review is done in-house
16mm B&W film prints	fair: have been circulated	Maintain in appropriate storage conditions	Keep until preservation actions taken on negatives	Storage is in-house

Table 2: An example of preservation strategy.

The strategy is just a simple table. The mapping of the archive gave us the breakdown into the different kinds of 16mm film in our collection (and the numbers, not shown in the table) – and the triage into issues of preservation urgency and need for access allow the material to be prioritised in a simple red-amber-green fashion. I'm pleased to report that we've since dealt with the vinegar syndrome and have eliminated the problem (for magnetic sound tracks on acetate – we don't have

any anymore!). We have finished the first three rows in the table, and hope to start pilot work this year on technology and a process for digitisation of the material in the fourth row: 16mm negatives.

A preservation plan is often very simple, once all the uncertainties around making a collection strategy and a preservation strategy have been resolved. Here is the BBC's preservation plan for the same collection of 16mm film:

**Preservation Plan: BBC 16mm film**

Type of material	Preservation Action	Service Provider	Batching	Outcome	Quality Control
16m mag sound track - masters	Digitisation at CD quality; 44.1 kHz sampling @ 16 bits; synch pulses recorded on 2nd CD channel	Three outside contractors selected by competitive tender	Monthly basis	One audio CD and one BWF file (on CD-ROM) per original mag sound track	Internal spot checking of each CD. Selective end-to-end checking. Done in-house.
16m mag sound track - duplicates	None				
16mm Ektachrome	Conservation for 2 more years; 10° C; 35% rh	In House			
16mm B&W film negatives	Conservation for 5 more years; 10° C; 40% rh	In House			
16mm B&W film prints	Conservation for 5 more years; 17° C; 35% rh	In House			

Table 3: An example of preservation plan.

This is a finite-duration plan. It refers to conserving the Ektachrome for two years – because we planned to concentrate on the urgent vinegar syndrome issue for the first two years. Then a new plan comes into effect, where we would list digitisation of the Ektachrome (and probably a review of the status of the B&W negatives).

With a plan, the column ‘preservation action’ contains entries that can be turned into costed business cases, and then into funded and hopefully very successful projects. The business case (or grant proposal, or funding request, or finance case) is often seen as difficult – but when the need is set within an overall strategy, and the work is focused on a specific action (one box in the above table), the problems should go away. The rationale is there, the action is specific, good cost estimates can be made – and the case goes through!

### Preservation principles and processes

There is still the question of ‘what, exactly, should we do (with a particular part of a collection)?’ Again, specific decisions are easier if they are made within a framework. Regarding audio and video, the overall ‘what to do’ has one general answer; *digitise, and start now*. There are no analogue options for maintaining access to analogue audio and video (again, film has a different answer). So the practical question comes down to cost-effective digitisation.

In general, digitisation is most cost-effective with a *factory approach* – an efficient workflow, a sort of assembly line or

mass production. The efficiency doesn’t require every item to be the same – because they won’t be anyway. Some materials will have problems and others won’t. The efficiency is through *division of labour* for the staff doing the work. If one person gets material from the shelves, another deals with metadata, and an audio or video technical expert is allowed to concentrate on audio and video (and not shelves and metadata), there are well-known efficiencies to be realised.

If the collection or the staff are too small for division of labour, it may be that an outside service provider is the best option: a company with all the needed equipment and expertise, and a business that depends upon efficient workflow.

There are also principles about what to get out of digitisation. The ideal is uncompressed video or audio. For audio, this ideal is easy to achieve, and indeed all the IASA-TC04 expert recommendations<sup>16</sup> are very clear that compressed audio is unacceptable for preservation. The minimum standard for audio is 16 bit quantisation at 48k samples per second. In the BBC we use the CD sampling rate of 44.1, but we see that as a small compromise made for practical reasons (to agree with what BBC radio uses in production). The material is stored as WAV files, and ideally as Broadcast Wave Files<sup>17</sup>. Everything else that a person would need to know about audio preservation and digitisation is in the IASA TC-04 document.

16. IASA-TC04, *Guidelines and the Production and Preservation of Digital Audio Objects*, 2nd Ed (2009)

17. [http://en.wikipedia.org/wiki/Broadcast\\_Wave\\_Format](http://en.wikipedia.org/wiki/Broadcast_Wave_Format)

Ingest Format	Migration format	Notes
<b>Low quality media</b>		
VHS tape	DVD	<b>Access</b> Perfectly adequate for VHS playback
VHS tape	MPEG-4 files	<b>Access</b> Adequate for quality. Minimum data rates (MPEG-4): 500k b/s. There are MANY potential access formats, and they come and go.
VHS tape	DV files	<b>Archive</b> (temporary) 25 M b/s, 12 GB/hr. Migrate to lossless for preservation.
‘low end’ digital files	Save as is, AND save as DV or lossless	<b>Archive</b> (temporary) Before format or DV format becomes obsolete, migrate to lossless for preservation.
DVD	DV files	<b>Archive</b> (temporary) 25 M b/s, 12 GB/hr. Migrate to lossless for preservation.
<b>Medium quality</b>		
U-Matic	DVD	<b>Access</b> Reduces quality; suitable only for viewing.
U-Matic	DV files	<b>Archive</b> (temporary) 25 M b/s, 12 GB/hr. Migrate to lossless for preservation.
DV, DVCAM	DV files (meaning .avi files with native coding)	<b>Archive</b> (temporary) transfers to computer at 25 M b/s, resulting in an .avi file ‘clone’ of the original DV tape. Migrate to lossless for preservation.
<b>High Quality</b>		
BetaSP, Digibeta, other pro formats	Uncompressed	<b>Archive</b> Uncompressed standard definition video: 200 Mb/s. About 100 GB for one hour, meaning 25 DVD-ROMs (or part of one data tape or hard drive).
BetaSP, Digibeta, other pro formats	Motion JPEG 2000 (lossless version)	<b>Archive</b> lossless compression, with a resultant data rate of around 90M b/s. About 40 GB for one hour.
DVCPRO50	.avi files, DV coding	<b>Archive</b> (temporary) As for DV, but at twice the data rate. Less susceptible to loss on future migrations. Migrate to lossless for preservation.
‘High end’ digital files	Save as is	<b>Archive</b> (temporary) Before format becomes obsolete, migrate to lossless for preservation.

Table 4: Video Migration Roadmap.

For video, there are problems:

- uncompressed video requires about 150 times more storage space, per hour, than uncompressed audio. So an affordable option for audio becomes a difficult decision for video;
- the material coming into the archive may already be compressed, with no possibility of access to an uncompressed signal;
- there is no simple, standard video file format – no direct counterpart to the WAV files for audio.

Professional broadcast archives can use the MXF<sup>18</sup> file format (a wrapper or container format). Other considerations include whether the archive's IT infrastructure prefers Apple or Microsoft file formats, and also what file formats are supported by any special equipment, such as video edit stations or software.

A final general principle is: if forced into use of a compressed format (by constraints at acquisition), do not move from one compressed format to another. The digital world has promised us perfect transfers – clones – every time we make a copy, without the *generation losses* inherent in an analogue transfer. But if we migrate from one lossy format to a different lossy format, there is an decoding and recoding process involved which can never increase quality, and is most likely to produce a drop in quality – a new, digital *generation loss*.

There is an argument for a partial exception, and that argument is inherent in the table 4 (p. 16). When going from low quality originals, one could move to a lossy compression format at a professional quality. One should never 'move sideways' from one compressed format to another of different encoding but similar datarate – that would be a generation loss. But one could climb the stairway to the heaven of uncompressed video by taking one intermediate step, 'resting' at a high-quality lossy format before moving to uncompressed a decade later. Full explanation of the table is giving in the PrestoSpace<sup>19</sup> wiki. The items marked "Archive (temporary)" are those which are moving to uncompressed video, but via one stage of use of a small amount of lossy compression.



Figure 4. Stairway to uncompressed heaven - by always going higher (but never going sideways). [http://www.flickr.com/photos/m\\_e\\_l\\_o\\_d\\_y/1002586303/](http://www.flickr.com/photos/m_e_l_o_d_y/1002586303/) Creative Commons licence.

### Stayin' Alive: Digital repositories and digital storage

All the strategy and planning and financial estimation and project management and digitisation finally results in a substantial body of audiovisual content existing as files, on some sort of mass storage. This is a major achievement:

- the content has been rescued from obsolete, fragile and decaying carriers;
- the connection between content (a Caruso recording) and carrier (a 78 rpm shellac disc) has been broken – forever; files can sit on any kind of digital storage;
- the content has been freed from shelves, and made ready for straightforward conversion to low-resolution version suitable for Internet access.

One might be tempted to sit back and bask in the rewards of a job well done: audiovisual content now firmly in the digital world. The question is: how firm is the digital world.

18. <http://en.wikipedia.org/wiki/MXF>

19. Migration Paths for Video Media: <http://wiki.prestospace.org/pmwiki.php?n=Main.Roadmap>

Unfortunately, digital storage devices have their own obsolescence, and on a three to five year lifetime rather than the 20 (or more) year lifetime of most analogue formats. File formats have their own obsolescence as well, especially the compressed formats used for web access. Proprietary formats such as Real Audio become displaced by more open (though not completely open) formats such as Quicktime and MPEG. New developments such as YouTube shift people's preferences for viewing formats.

The result is that requirements for vigilance and management – for continuous maintenance of a collection – remain as necessary for digital content as for analogue. The managerial requirement is the same as ever, but there are new problems and new coping strategies in the brave new digital world.

As an example, the BBC transferred audio material from 1/4" (6mm) tape to broadcast wave files, and stored those files on writeable DVDs (DVD-ROM). That format was chosen as an interim measure, as there was no general BBC-wide mass storage system at the time. A few years later, and this material was transferred to files on USB-connected terabyte hard drives (30 of them, holding 40 thousand hours of archive content). The good news: that migration had 1% of the cost of the original digitisation, because it was digital-to-digital and could be automated to a considerable extent. The bad news: these hard drives have had various problems, and a professional-level BBC-wide mass storage system does now exist, so the material is being migrated again. Also, at the same time as transfer from DVD-ROM to hard drive, a back-up copy was made on LTO-3 datatape. This format is about to be replaced by LTO-4, and by the time that LTO-5 comes on the market, the LTO-3 tapes will have to be migrated.

There are many such examples. The Austrian Mediatheque<sup>20</sup> is on its third datatape robot since the year 2000, as are the Swedish broadcaster SVT and the French national archive INA. Even robots struggle to survive in the digital world! Another case is material 'born digital' on formats such as DAT and minidisc, where again the formats are obsolete and the content requires migration.



Figure 5. Stayin' Alive: robot survival in a digital world. Robot: <http://www.flickr.com/photos/punkjr/518165822/> Creative Commons licence.

Fortunately there are sources of aid<sup>21</sup>. Many areas of modern life are creating digital content, from word processing and book scanning to scientific and business data. A whole technology has grown up, centred around digital repositories and digital preservation standards and technology. Most of this technology has been developed by the 'conventional' library world, as it moved to electronic documents. The key issue for audiovisual content is to find ways to use digital library tools and processes on audiovisual files. The new project PrestoPRIME is addressing that problem.

20. [www.mediathek.ac.at/](http://www.mediathek.ac.at/)

21. A brief summary is Richard Wright, *Preservation of Digital Audiovisual Content* (2008); [www.digitalpreservationeurope.eu/publications/briefs/audiovisual\\_v3.pdf](http://www.digitalpreservationeurope.eu/publications/briefs/audiovisual_v3.pdf)