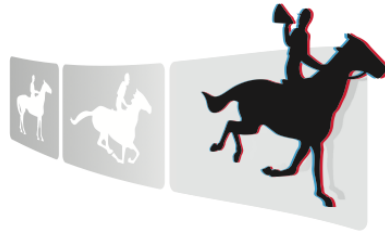


KEEPING AUDIOVISUAL CONTENT ALIVE



PRESTO
CENTRE

Tutorial: Planning your preservation project



© Richard Wright and PrestoCentre 2011

Published by PrestoCentre, <http://prestocentre.eu>

The editors and authors of this work assert their moral rights to be identified as such.
This work may be copied, distributed and transmitted for all but commercial purposes. You may not alter, transform, or build upon this work. For any reuse or distribution, you must make clear to others the license terms of this work, including any embedded licensing metadata.
Enquiries concerning use of this work outside these terms should be sent to PrestoCentre, % Netherlands Institute for Sound and Vision, PO Box 1060, 1200 BB Hilversum, The Netherlands.

This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/> or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

© Cover illustration copyright by the Netherlands Institute for Sound and Vision

Introduction

Archives tend to take a two level approach to preservation planning:

- Firstly, they define their high-level preservation strategy. This defines what needs to be done, when it needs to be done by and how much it will cost in total.
- Secondly, they consider the operational aspects of preservation, by focusing on the details of how preservation will be done on a day-to-day basis. This is the actual preservation plan.

The strategy may say 'conserve another 10 years and then review', or it may say 'make new master copy'. The strategy may even say 'make new digital master copy'.

But it is the plan that defines what the conservation method will be. The real complication comes in deciding what exact method will be used.

Preservation projects are invariably a balancing act between:

- Cost: how much the preservation project will cost
- Quality: what quality can be achieved
- Time: how long the project will take
- Volume: what volumes of material can be preserved within that time

These are the four key factors of preservation. Making trade-offs between these factors is an engineering approach to preservation and aims to find the best compromise given the circumstances. This is sometimes referred to as the 'cost of quality' approach. In this way, the usual trade-offs between 'better', 'faster' and 'cheaper' can be explored.

1. Building upon the preservation strategy

As said, the preservation plan is a detailed and concrete plan for following your preservation strategy.

If, in your preservation strategy, the decision has been made to preserve the collection by transferring objects to a new carrier, the details about the new carrier need to be specified in your preservation plan.

A vital preservation decision is: what to migrate onto (what new format to use). The basic decision about a new master copy is whether to make an analogue or a digital copy. PrestoCentre advice is that, except for film, there is no reason to ever make an analogue copy of audio or video. It will cost more -- and give less benefit -- than digitising.

In consequence, the main new information in a preservation plan is:

- the exact specification of the digital object that will be made
- who will make the digital object
- how long the process will take

For a large project, there is a lot more detail in a plan -- because a large collection cannot send all its material out for transfers at one time. So a large collection breaks the digitisation into phases, and the plan should define the phases. For instance, a collection of U-Matic tapes covering a 10-year age range might be done in groups of one year's worth of tapes at a time, working from the oldest to the youngest.

Prioritisation requires rules to be defined by analysing both the business and technical requirements. The business and technical factors cannot be treated independently. For example, one part of the archive might have a high business value, but actually be given a low priority for digitisation

since the carriers are not in danger of decay and there are still playback machines in production so there is not a risk of losing access. On the other hand, another part of the archive might have moderate business value, but the carriers are rapidly decaying and hence transfer is urgent to avoid widespread loss.

Priority is typically governed by properties such as content, genre, usage, carrier type and condition.

Here is the (hypothetical) strategy for the BBC 16mm film collection, followed by a simple preservation plan:

Preservation Strategy: BBC 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

Preservation Plan: BBC 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			

Of course there are a lot more decisions to be made for actual running of the transfers.

It is easy to write 'batching: monthly basis'. In an actual project, the entire collection of magnetic sound tracks has to be identified (in the catalogue if possible), identified on the shelves, segregated immediately to a new room (because of the threat of contamination), and then a method has to be devised for sending the right amounts of material each month to each contractor. The checking has to be decided upon and organised. Service level agreements with the contractor are needed – and these agreements have to be monitored and managed.

A lot of detail is about metadata: how the items are to be found, and how the 'outcome' items are to be labelled and identified. One of the easiest ways to save time and money in any large project is to use bar codes for identification of items. Trained audio and video specialists can get on with their specialised work in an efficient manner if bar codes are used. Ideally every physical new item will have an integral bar code and packaging with a bar code – and the bar codes will agree with the catalogue for the collection.

If the metadata and physical identification is not thought through and automated, up to half of a trained specialist's time can be expended on purely logistical issues of identifying and re-labelling of items. Quality control also suffers with manual identification, as labels can get mixed up, words get misspelled and a range of other human errors can creep into the process.

2. Budgeting your preservation plan

When budgeting your preservation plan, it is important to take a Total Cost of Ownership (TCO) approach where all cost aspects are considered (at least to start with, some aspects might be discarded later if inappropriate or negligible).

The 'total cost of ownership' includes:

- production costs of initially digitising content in an archive;
- operational costs of maintaining, supporting and managing the archive;
- unavailability costs, such as lost business, or the need for additional support should the content or services of an archive become unavailable for some reason.

When calculating production costs for preservation transfers, there are generally two approaches, which can be combined:

- **Predict the cost**

This is done by defining the preservation processes in detail for a particular archive based on content, carriers and condition; costing each step of the process using estimates or quotes; and aggregating the individual costs to create an estimated total cost. PrestoCentre has a variety of tools to do this, as well as reports on costs from different projects.

- **Measure cost through pilot studies**

The best way to estimate costs is to do a pilot project, and work out the average time taken for two cases:

- a simple transfer, where nothing goes wrong
- a problem transfer, where extra steps are needed to make the transfer work

It is then up to the user of this model to convert from time to cost, and use the model to estimate full project costs.

A representative sample of archive content is selected and then a small-scale preservation project is executed using this sample. External service providers or existing infrastructure is typically used to minimise large capital investment. The benefit of the pilot study approach is that it will typically better reflect the true state of items in the archive and the realities of running a preservation project in practice.

PrestoCentre has a detailed model for estimating costs of a transfer project. At the heart of the model is a per-item or per-hour contractor cost. Although the model has a default value, the user really should update this value, based on negotiation with a service provider.

If transfers are being done in-house, there is no service provider to ask. However cost estimation for in-house projects are filled with uncertainties, because of the many different ways that collections count their costs.

The other key item of information required by the model is an estimate of the percentage of the material that will have a problem. Typically problem material costs something like four times as much (takes four times as long) as non-problem material. This ratio means that for a project with 20% problem material, half the budget goes on the problem 20%, and the other half goes on the simple 80%. If money is insufficient to pay for everything that a collection needs (and all archives have budget problems), then it makes sense to get 80% of the archive transferred rather than 20%, by concentrating on the straightforward material.

This is the principal of triage. If resources are limited, triage is used to get the best result from the limited resources. For archives, resources are always limited, and triage is always preferable to simply allowing chance to determine what is saved and what isn't.

Triage example:

Consider transfer of 1000 items

80% with no playback problems

BUT - problem items take 4x as long to process

Consider a cost per non-problem item of 100 and a budget of 100 000

To these production costs, the operational and unavailability costs need to be added in order to have a full picture of the costs concerned with preservation.

3. Changing circumstances

When planning a preservation project that might take years, it is obviously important to think about how circumstances might change over time. An archive is unlikely to be in the same position in five years time as it is in today, let alone ten.

From a business perspective, changes over time can be hard to predict, for example:

- An archive might acquire one or more new collections, which in turn increases the demand for preservation.
- New funding initiatives might arise and allow much more to be achieved in a preservation project than was originally planned. The reverse could happen and sources of investment into the archive might dry up.
- The user needs for archive content can change over time, especially in broadcasting, if they depend on what's 'fashionable' at the time.
- Public perception and government policy can change, for example resulting in new initiatives for increased archive accessibility for public benefit.
- New technologies might open up new distribution channels and markets and hence change the user base of the archive.

strategy	Regular items	Problem items	Result
Strategy 1: random	500 @ 100 each = 50 000	125 @ 400 each = 50 000	62.5% of job complete
Strategy 2: easiest first	All 800 easy items transferred = 80 000	20 000 left to do 50 hard items	85% of job completed
Strategy 3: hardest first	All 200 hard items transferred = 80 000	20 000 left to do 200 easy items	40% of job complete
Strategy 4: partnership	Service Provider does 800 easy items = 80 000	Archive "saves" 20 000 and handles the 200 hard items themselves	100% of job complete and project completes under budget - if internal work is counted as effectively free

For all these reasons, it is important to have a flexible preservation plan that can be revisited and revised on a yearly basis to allow change to be managed. The plan should take account of time dependent factors as best as possible. These include:

- Technical obsolescence
- Content degradation
- Inflation and deflation

These are described in more detail in the sections below.

3.1 Technical Obsolescence and Storage Capacities

Technical obsolescence and rapid advances in storage technologies (size, cost, capacity, management etc.) mean that the digital archive is not a static entity based on one solution, but typically a continuously changing entity that consists of a mixed set of storage solutions and media types.

The formats, media, hardware and software that are commonplace in archives today are unlikely to be that way forever. Equipment and media formats become obsolete and unsupported, and the associated operators and skills die out (sometimes quite literally!).

No formats or equipment are immune from this problem – the question is whether it will happen to you (or to your service providers) in the lifetime of your preservation project. While technical obsolescence is influenced by discrete events, for example a manufacturer discontinuing a particular playback device, in general technical obsolescence is a continuous process and there isn't a specific point where transfer of a particular carrier type suddenly becomes impossible.

Technical obsolescence applies to digital media and storage solutions as well as to analogue material. Indeed, technical obsolescence in the digital

world can happen even faster than in the analogue world.

Even if obsolescence is no problem for the digital media you use, a migration plan might be beneficial anyway because of growing performance and capacity of newer media.

A specific mass storage system (e.g. a particular tape robot and tapes) will become obsolete roughly once every 15 years. As we have seen in the past years we know that for example the LTO consortium brings a new technology (with double capacity) every 2-4 years. Backwards compatibility is guaranteed within two technologies, at least for reading. Migrating your collection every 7-10 years will increase your storage performance (due to higher read and write speeds with new drives) and capacity (migrating every 8 years involves skipping at least 2 technologies and quadrupling the storage capacity).

3.2 Content Degradation

If an archive's content didn't decay, then one preservation strategy would simply be to just keep items on the shelves until they are either no longer needed, or until transfer becomes so cheap that preservation can be done at a greatly reduced cost.

Most media decays, and if not transferred sooner rather than later then it will often be lost forever. In general, you can slow decay down (with low temperature and low humidity storage), but you typically can't stop it completely. Other factors include wear and tear and accidental loss.

Predicting the timescales of a preservation project based on decay over time can be very important when justification will be in terms of 'avoidance of loss'. The deadlines imposed by decay (and technical obsolescence) can mean that substantial and rapid action is often needed to avoid irrevocable loss. This is especially true when dealing with a large backlog of work (e.g. large existing analogue holdings) as opposed to steady-state needs (new acquisitions of analogue material).

As with technical obsolescence, it should be noted that degradation due to carrier decay or wear and tear is a continuous process. There is also typically a 'spread' to degradation, i.e. some items will decay faster or slower than the average, even when stored in identical conditions.

3.3 Financial Changes Over Time

Labour costs are typically subject to inflation, i.e. they increase over time. This doesn't mean that changes are as smooth and continuous as they might be for more general inflationary measures such as the consumer price index. For example the cost of skilled operators for a particular format will be dependent on the 'lifecycle' of that format, especially when the format becomes obsolete.

On the other hand, storage costs are typically subject to deflation, i.e. the cost falls over time. For example, the cost of mass storage, whether for hard discs or for data tape, has been dropping sharply (reducing by 50% every 18 to 24 months) and that trend has continued for forty years, so it should continue for at least another ten and probably another twenty years.

This means that storage should be bought as late as possible, and it makes sense to buy storage in small rather than large quantities, and only when needed. It is pointless getting a 20% quantity discount now, if a 50% reduction is available just by waiting 18 months.

The essential condition for benefiting from these cost reductions is that technology bought in bits and pieces rather than all at once has to fit together.

4. Funding your preservation plan

Given the overall business environment that a digitisation project must operate within, it can at times be rather a daunting undertaking to make the assessments needed to define your plan.

The whole issue of trying to attach quantifiable value to archive holdings is a difficult one. It is far more common for archives to talk in terms of the 'cost of loss' if they could no longer access their holdings, than it is for archives to predict the positive return on investment from a preservation project that maintains or improves access.

It may be possible to identify some specific items that have known high value by assessing the programme material in the archive, for example major historical events or a popular TV series. But what about the rest of the archive? This is usually the majority of the content!

The general message around funding is that "access attracts support" – but it is up to the individual collection to turn that generalisation into something specifically useful – and profitable – for themselves.

If there are commercial possibilities for a collection, conventional thinking was that you shouldn't give away anything that you can sell. But even conventional thinking would agree that you have a better chance of selling something if you have a widely distributed catalogue – with nice pictures. Putting archive material on the web should be thought of as a catalogue, not as a lost sale. There are now statistics from collections which have 'given away' web versions of their content, showing dramatic increases in business resulting from the simple fact that the free material on the web was effective advertising.

Conclusion

After going through these four steps you have created a basic preservation plan to work with in the future. To add more detail to this plan, research the different steps in depth on the Presto-Centre website.

KEEPING AUDIOVISUAL CONTENT ALIVE



PRESTO
CENTRE



SEVENTH FRAMEWORK
PROGRAMME