



Deliverable D22.8 (2007) Preservation Status

Annual Report on Preservation Issues for European Audiovisual Collections (2007)

DOCUMENT IDENTIFIER PS_WP22_BBC_D22.4_Preservation Status_2007

DATE 02/03/2007

ABSTRACT This is the third annual report on the preservation status of European audiovisual material. The report gives information on the cost of digitisation, and the cost of the 'digital preservation' of the results of digitisation. The report also reviews current information on the value of audiovisual content, and concludes with updates on archive access and on European political and commercial developments.

KEYWORDS Preservation Audiovisual Archives Status Training Europe Value Cost Access

WORKPACKAGE / TASK WP22

AUTHOR, COMPANY Richard Wright, BBC

INTERNAL REVIEWERS David Jordan, BBC

DOCUMENT HISTORY

Release	Date	Reason of change	Status	Distribution
1.0	22.02.2007	First Draft	Living	Confidential
2.0	27.02.2007	Final circulated for review	Living	Confidential
3.0	02.03.2007	Final submitted to prestospace	Living	Confidential
4.0	02.03.2007	Final – passed external review	Complete	Public

Contents Table

1. Document Scope	3
2. Executive Summary	3
3. Cost and Value Overview	4
4. Costs: Digitisation and Associated Activities	7
4.1. Short projects: two years or less	7
4.2. Long projects: more than two years	8
5. Digital Preservation	14
5.1. Digitisation - solution or problem?	14
5.2. Maintaining access.....	14
5.3. Media failure and system obsolescence	17
5.4. Where to learn more	18
6. Digitisation Costs in Europe	19
6.1. Cost of Digitisation	19
6.2. Cost of Digital Preservation.....	21
7. Value of European Audiovisual Heritage	27
7.1. Commercial and Public value.....	27
7.2. Releasing value.....	30
7.3. Interactions.....	32
7.4. A call to arms	34
8. Access	36
8.1. Public Access and User-Contributed Content	36
8.2. Archive Access.....	36
8.3. Commercial Access	37
8.4. Research.....	38
9. European and Global Support for Audiovisual Preservation	39
9.1. EC support for audiovisual preservation	39
9.2. TEL and i2010 Digital Libraries and EDL	39
9.3. Google and Euro-Google	40
9.4. The European Archive	40
9.5. The Global Dimension.....	41
10. Service Industry Developments	42
10.1. Vienna Workshop	42
10.2. Storage Service Provider offerings:	42
11. References	44

1. Document Scope

This report is an annual public document, aimed at persons responsible for audiovisual collections and giving a status report on audiovisual preservation issues across all EC countries. The report gives information on the cost of digitisation, and the cost of the 'digital preservation' of the results of digitisation. The report also reviews current information on the value of audiovisual content, and concludes with updates on archive access and on European political and commercial developments.

2. Executive Summary

Two years ago, we reported on data from 21 countries and direct contact with 31 collections – and found 20 million items (around 10 million hours). That report concentrated on the physical state of the items, plus some data on preservation projects and funding.

Last year, the TAPE survey in 2005 had responses from nearly 400 archives, and found about 25 million hours of film, video and audio – of which an estimated five million hours was already identified in last year's PrestoSpace survey.

The TAPE survey had results that are even more significant than this additional 20 million hours of identified material: it showed the low *state of awareness* of preservation as a problem, and the even lower *state of preparation* for meeting that problem: **half the archives had no controlled storage conditions, half had no regular equipment maintenance, and 2/3 had no systematic preservation programme.**

This year we look at the cost of digitisation, and then at the long-term costs for 'digital preservation' *after* digitisation. We give the costs, now, for digital storage – covering a range of storage options – and we forecast the trend for costs over the next twenty years. The result is sufficient information for archives to create 20-year plans for their new, digital archives. These plans are essential because digital materials, just like analogue ones, will not last forever and are just as much in need of *maintenance* as were the old 'tapes on shelves'. In fact, timescales for obsolescence of digital technology are, if anything, shorter than for the old materials.

To balance the cost issues, we also summarise current information on value: what it is and how to calculate it. *Public value* is especially relevant to heritage and broadcast collections, and there is growing awareness of the fact that the public value of audiovisual materials far exceeds previous narrow concepts of commercial value.

Finally, the report provides an update on developments in online access, and new initiatives relevant to audiovisual materials – both political and from the audiovisual services (facilities) industry.

3. Cost and Value Overview

The most common question about archive preservation is: how much will it cost? The answer is usually complicated, because the simple answers are usually simple only because they leave out all sorts of important costs, and so are underestimates.

Underestimates help nobody. They may be seen as making it easier to get approval, but will inevitably lead to cost over-runs, which either bring the project into disrepute, or force the project to take short-cuts (which always compromise quality) or bring the project to an end before the job is done – because the funds have run out. Many projects experience all these problems.

The best thing the Presto project ever did, in my view, was to collect data representing for projects where there was relatively complete cost information, and publish those. These costs were considered high, which means they have allowed other projects to do better – which is always a good thing. These high costs figures have also prepared archivist and managers for the realities of preservation projects – another good thing.

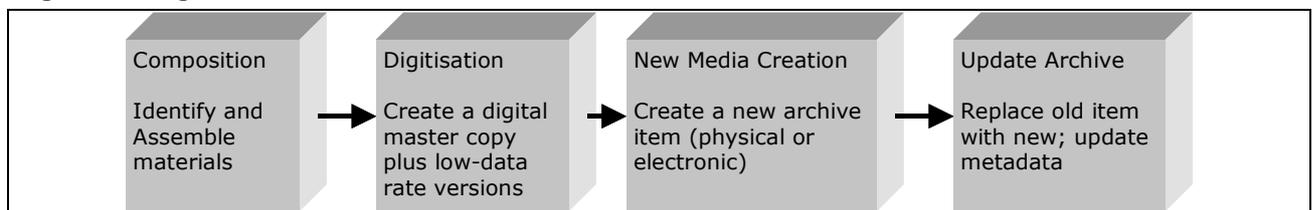
Presto published 'cost per hour of material' figures [[see table]], which is a good standard number, but it is not the full picture. The Presto figures covered a digitisation transfer, not everything associated with preservation. The figures included transport and equipment and operators and handling and indeed everything we could track down that was directly associated with a 'preservation transfer' project – but there are many other potential costs.

Table 1- Comparison of costs and usage between traditional and digital systems¹

	Digitisation cost, per hour
Film	2000
Video	200
Audio	120

Presto also published a model (see reference 2) for a range of activities associated with digitisation, from all the actions needed to identify materials for transfer, gather them together and transport them, and their metadata, to the digitisation area – to providing web access and updating the metadata (catalogue). These additional activities are shown in the following figure.

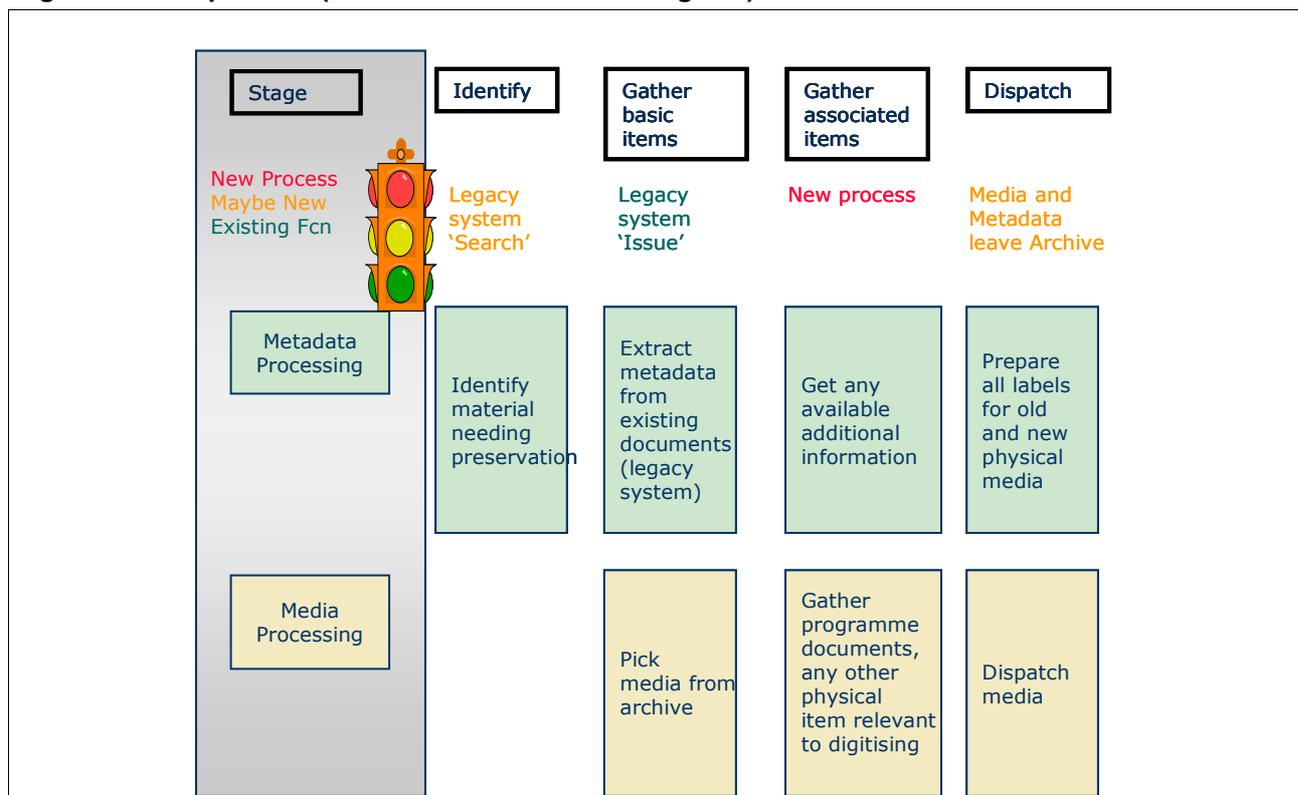
Figure 1 – Digitisation and associated activities²



Of course, the above figure is by no means everything. Each of the boxes could be broken down into many more steps – and costs. As an example, Figure 2 – Composition shows breaking that process into five steps, each with two parallel streams:

- actions concerning documentation (metadata);
- actions on the actual media.

Figure 2 – Composition (of a batch of materials to digitise)



[The interested reader is referred to reference 2 for more such detailed figures, covering *Digitisation*, *New Media Creation* and finally *Update Archive*.]

We review all the costs of digitisation – and of everything that goes with digitisation in a preservation project – in **Section 4** Costs: Digitisation and Associated Activities.

Digitisation is not all there is to preservation, because digitisation is just a step in the (hopefully) long life of an audiovisual item. After digitisation, the “bits in the bit bucket” need to be kept readable (*persistence* of the bits) and usable (*currency* of the file format). *Persistence* and *currency*³ are the essential requirements of *Digital Preservation*.

All forms of digital data have digital preservation issues. For audiovisual materials that have been digitised, digital preservation is just another aspect of the *maintenance*⁴ that is essential for the survival of any audiovisual item. PrestoSpace uses the term *maintenance* to cover all areas of ongoing preservation work, before or after or completely independent of digitisation. Use of the term *maintenance* is an attempt to stress that archives should expect, indeed demand, funding for *maintaining the viability* of archive content – and that the funding should be part of every annual budget, not just something special for performing a rescue operation when matters have reached a crisis.

Digital preservation is presented here as a form of maintenance, but digital preservation has become a distinct area of study and expertise. The good news for audiovisual archivists is that because all forms of digital data have preservation requirements, there is a great deal of study and knowledge and technology in common, across all forms of digital

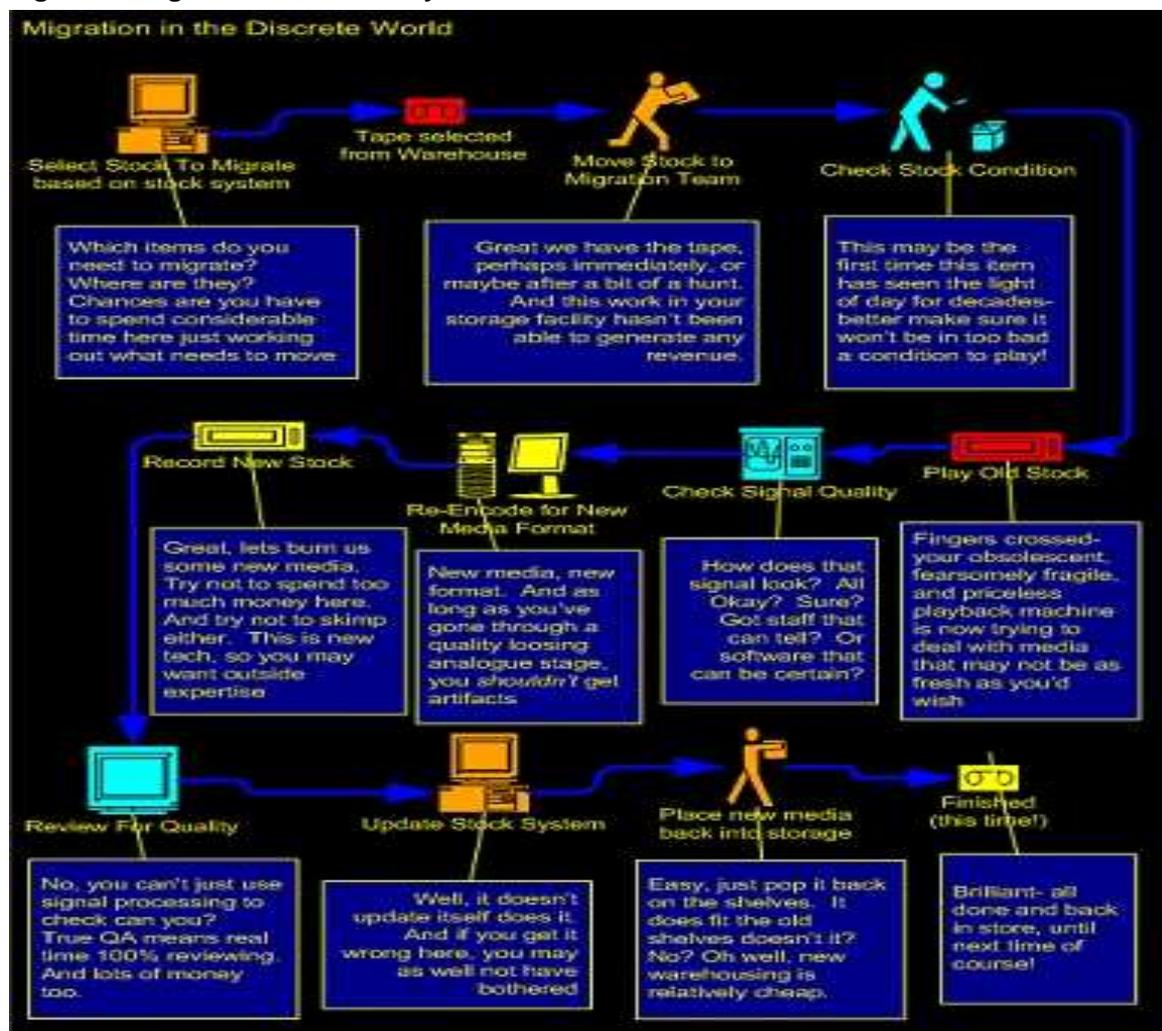
preservation. This whole area – digital preservation as applied to audiovisual material – is presented in **Section 5** Digital Preservation.

Some specific costs are then presented in **Section 6** Digitisation Costs in Europe.

This paper intends to present not just the cost of everything (audiovisual), but also the value of everything. Value is probably a more complex subject than cost, because it has to do as much with a philosophy of what should and should not be included in 'value judgements' – as it has to do with objective measures of economic activities. Fortunately there is a discipline that combines exactly these two fields – philosophy and metrics – and that field is economics.

The economists' view of value – in particular *public value* – is presented in **Section 7** Value of European Audiovisual Heritage.

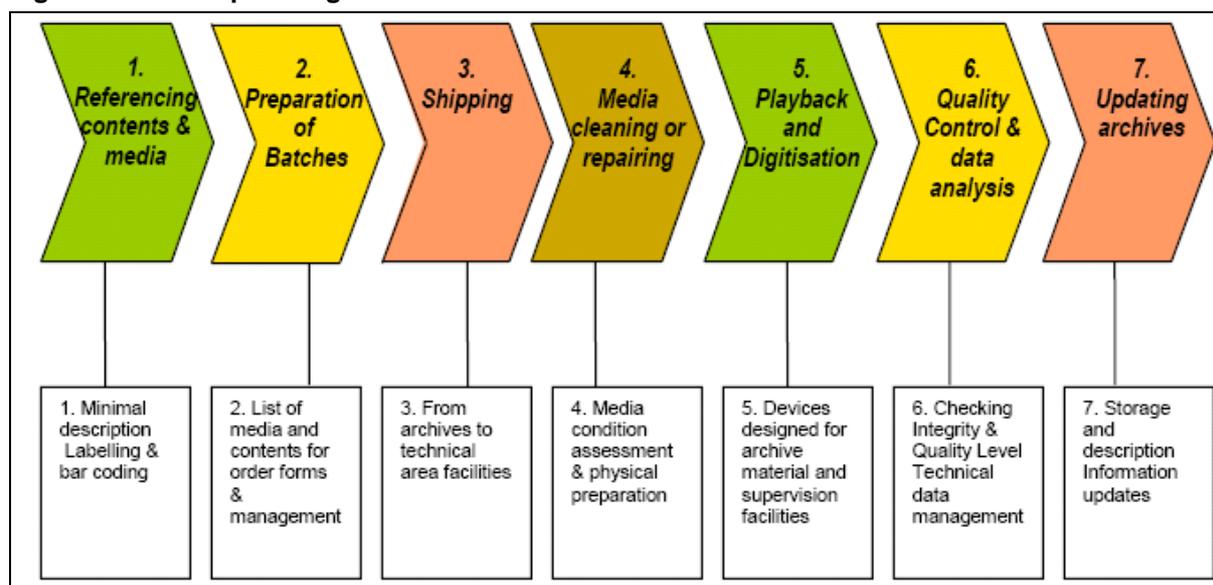
Figure 3 - Migration of 'hold it in your hand' media⁵



4. Costs: Digitisation and Associated Activities

Section 3 introduced the full range of cost and value issues, to be covered in this section and the next three. In this section, we just look at the processes around digitisation, according to the PrestoSpace model shown below in Figure 4. Immediately there is a major distinction to be drawn – between relatively short projects (roughly two years or less) and anything longer. The distinction is vital, because for short projects the costs can be considered constant, while for the longer projects all the costs really have to be considered as variable, changing over time in important ways.

Figure 4- PrestoSpace Digitisation Processes⁶



4.1. Short projects: two years or less

PrestoSpace has considerably advanced the clarity of thinking in the area of audiovisual digitisation and storage, and has a detailed cost model⁷ [[use DPC/DCC if nothing else]], which has been simplified and implemented as an online tool with a detailed online explanation⁸. The online tool supports basic preservation planning, and is especially useful for short projects with fixed costs. This model is explained in detail in a PowerPoint presentation available as part of the PrestoSpace Preservation Guide⁹. A great amount of other PrestoSpace material is relevant to effective digitisation. Indeed the whole project centres on that issue. Of particular significance is the general information on how to deal effectively with digitisation management issues as explained in Deliverable 5.3 Tools for automation of difficult media¹⁰. There is also a full description of how to approach a digitisation project, as part of the whole collection strategy, covered in the various sections of the online Preservation Guide¹¹.

Working out digitisation costs does not have to be complicated. Here are the essential cost categories:

- Number of items to digitise
- Media type and condition

- Cost of transfer per item
- Digitisation quality and usage
- Basic storage costs

Section 3, above, listed three other areas of activity (and cost) in addition to digitisation:

- Composition – everything needed to get material ready for digitisation
- New Media Creation – writing “the bits” onto something
- Update the Archive – the catalogue needs to know about the identity and location of new materials, and any change in location or status of the originals – and the physical materials, old and new, have to be gotten to their proper places.

The PrestoSpace simple cost model covers Digitisation and (digital) New Media Creation (and a bit more). Composition and Update the Archive are activities that have to happen, but may or may not be included in the financial planning for a project – especially a small project – because these are steps that are usually handled by existing archive staff.

Section 6 below gives actual costs – real numbers. For people who cannot wait, archives report their current video transfer costs as €131 per hour. Interestingly they expect their planned projects to cost considerably more (164€ per hour), which may be ‘the voice of experience’ that should be listened to.

The figure of “150€ per hour for digitisation” provides a simple answer to a simple question “how much does it cost”. However the PrestoSpace view is that “the cost of preservation” is not a simple question, and does not have a simple answer. The figure quoted is about as useful as, for instance, the average price of a car, any car, new or used, sold in the UK last year. Averaging over a wide range of situations does not necessarily reveal anything.

4.2. Long projects: more than two years

It is even harder to estimate costs for longer projects, because the costs of the individual *cost factors* change over time. These changes can be put in the following categories:

- Small increases – mainly “cost of living”, and predictable
- Big increases – various factors, often hard to predict
- Decreases – one specific cost does get smaller: digital storage

Small increases: the factors relating to staff costs will go up according to the general “cost of living”, which is largely predictable – certainly over roughly the next five years. Cost of premises: floor space, light, heat, electricity also go up by the same general amount. In the PrestoSpace model, the ‘cost per hour of a transfer’ should be increased by ‘cost of living’ for projects from three to five years in duration.

Projects longer than five years need to add an additional uncertainty or contingency factor, or plan a review and re-budgeting stage at no more than five years into the project.

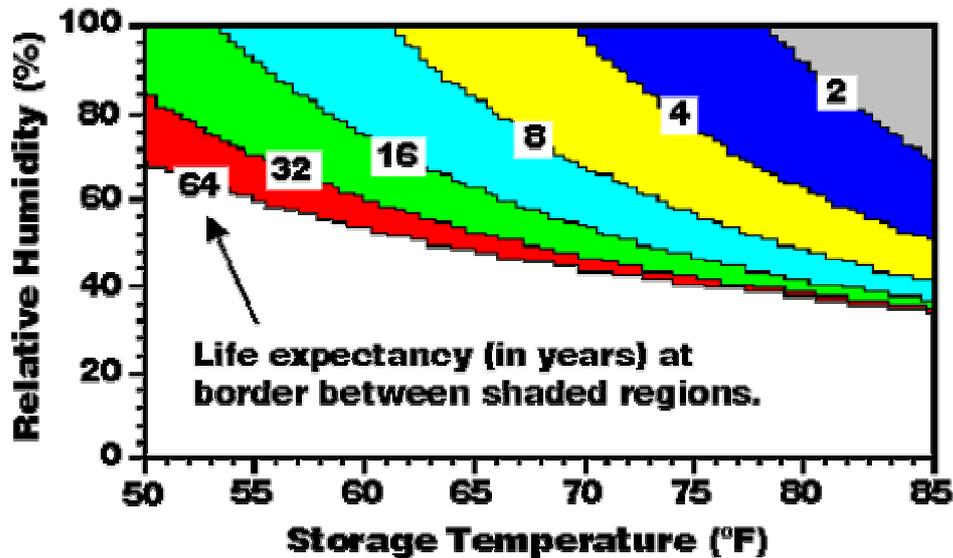
Big increases: The four obvious sources of large increases in cost are:

- Degradation of the original materials
- Availability of equipment
- Maintenance of equipment
- Availability of staff

Degradation: The PrestoSpace cost model, based on the *Preservation Factory* concept (see

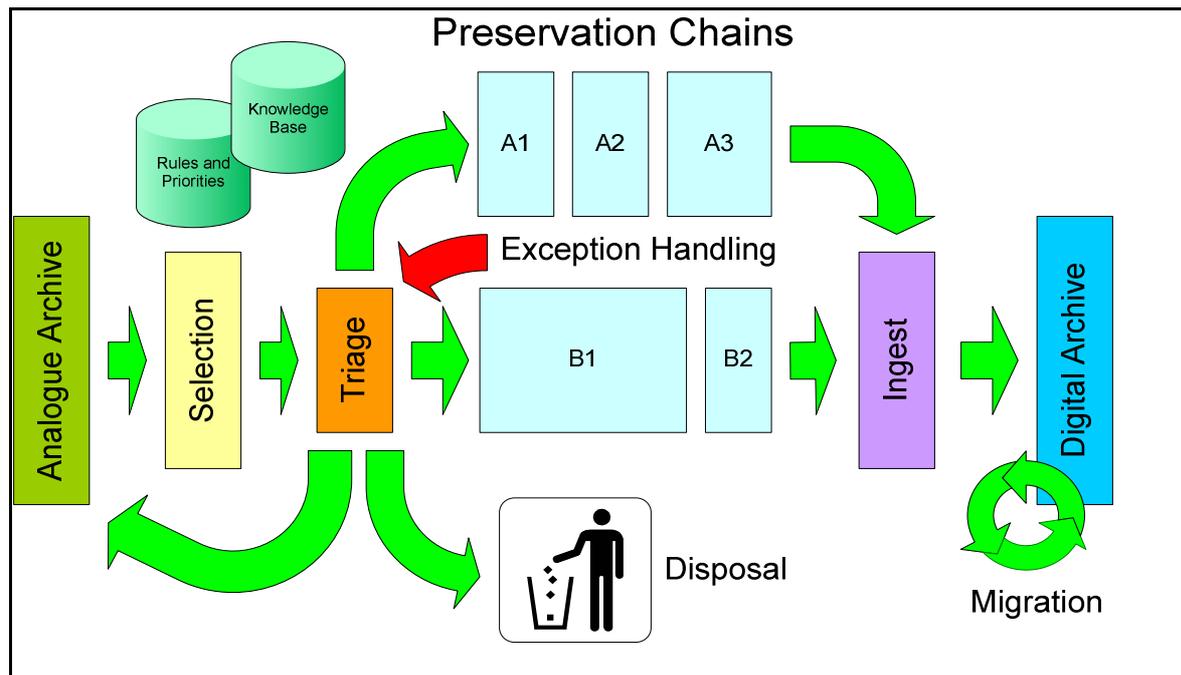
Figure 6, below) of maximising the efficiency of the workflow, assumes a certain proportion of the material will be problem-free, and go through the process at minimum cost. The remaining material will cause difficulties, and have to be taken out of the main workflow and given special treatment, such as cleaning and then re-try by a more skilled (and more expensive) operator.

Figure 5 - Tape Life Expectancy as a function of storage conditions¹²



As a general guide, problem material takes about four times as long to process. This means that a project with 20% problem material costs 60% more than a process with no problem material.¹³

The degradation of audiovisual materials over time is not a uniform process, and is also very hard to predict. It is a combination of the age and the entire storage and usage history of the material – plus an element of chance (representing the variance of the decay process). The chemical process is also non-linear, in that the degradation follows an exponential curve. Strictly speaking, the proportion of items in a collection affected by decay follows an exponential curve – which in simpler terms means that once materials start having playback problems, the problems will get worse – and will get worse faster and faster.

Figure 6- Preservation Factory with separate workflow B for 'problem material'¹⁴

Once 60% of tapes have playback problems, total cost of a digitisation project will escalate to %280 of the “problem-free cost”.

PrestoSpace has tried to include the exponential model of degradation in a full cost model. Unfortunately we have been unable to verify the model against known cases. The conclusion is that we are left viewing degradation as an unpredictable cost. For short projects, the level of problem tapes should be predictable from a pilot run of randomly selected samples. For longer projects, the PrestoSpace recommendation – and the standard practice for large preservation contracts – is that the percentage of problems tapes actually observed needs to be noted, every year, and any increase means the budget for the next year needs an equivalent increase.

There are various strategies for dealing with unpredictable costs on a fixed budget (all of them less than ideal):

- Limit work on problem tapes to a fixed time, and then stop
- Abandon problem tapes entirely, as soon as the problem occurs

In either case, the archive is left with a pile of difficult material at the end of a project. The most optimistic thing that can be said about that situation, is that at least a new ‘clear the Augean stables’ project could then be launched, in full knowledge of the percentage of problem tapes – because it would be 100%.

Availability of equipment: Equipment might be seen as increasing in costs according to inflation, or even decreasing (as is the case for home computers). Unfortunately digitisation relies on use of equipment that is no longer in production. As equipment becomes scarce, the cost could rise. In any case, cost is unpredictable – and the time needed to find equipment could have the biggest increase.

Maintenance of equipment: Maintenance could also be expected to follow cost of living, but the problem is the unavailability of parts – which means the cost of maintenance is exactly as problematic as is the case for the equipment itself.

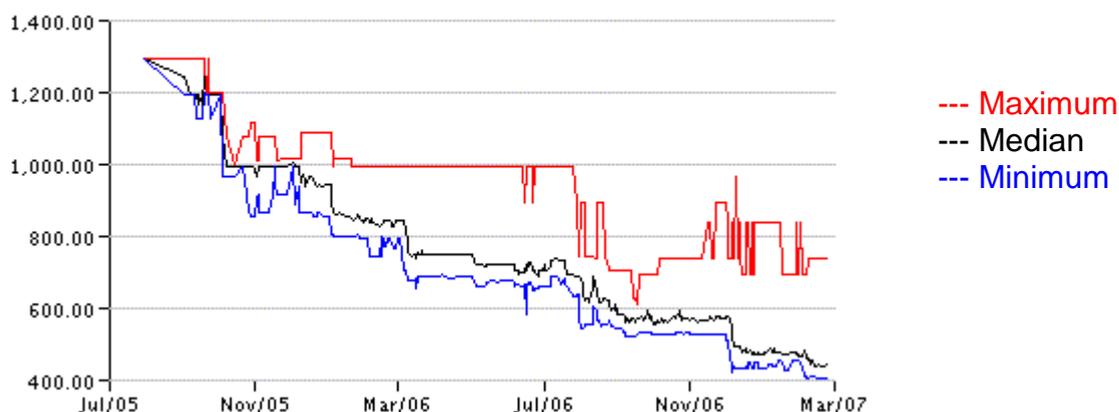
Availability of staff: It may be wishful thinking, but aged audiovisual engineers could become a premium commodity. As with equipment and parts, the really unpredictable aspect concerns whether or not experienced staff can be found.

All the problems, headaches and associated costs relating to degradation, equipment, maintenance and staff will only get worse over time – in an unpredictable fashion.

Decreases: cost of storage. The next two section will go into the complexities of digital storage, with special attention to which cost factors go down, and which do *not* go down, including storage cost factors which rise. But for storage *media* – which is not at all the whole story as Section 5 will show – the costs do go down, dramatically. Digital storage costs have been following Moore's Law¹⁵ for about 40 years, and can be expected to do so for the next 20. The 'law' (for complexity of integrated circuits) says that complexity doubles (for the same cost) every two years. This has been loosely reformulated by varies parties as "Moore's Law for Storage" and claiming a similar doubling of digital storage capacity (for the same cost) every two years.

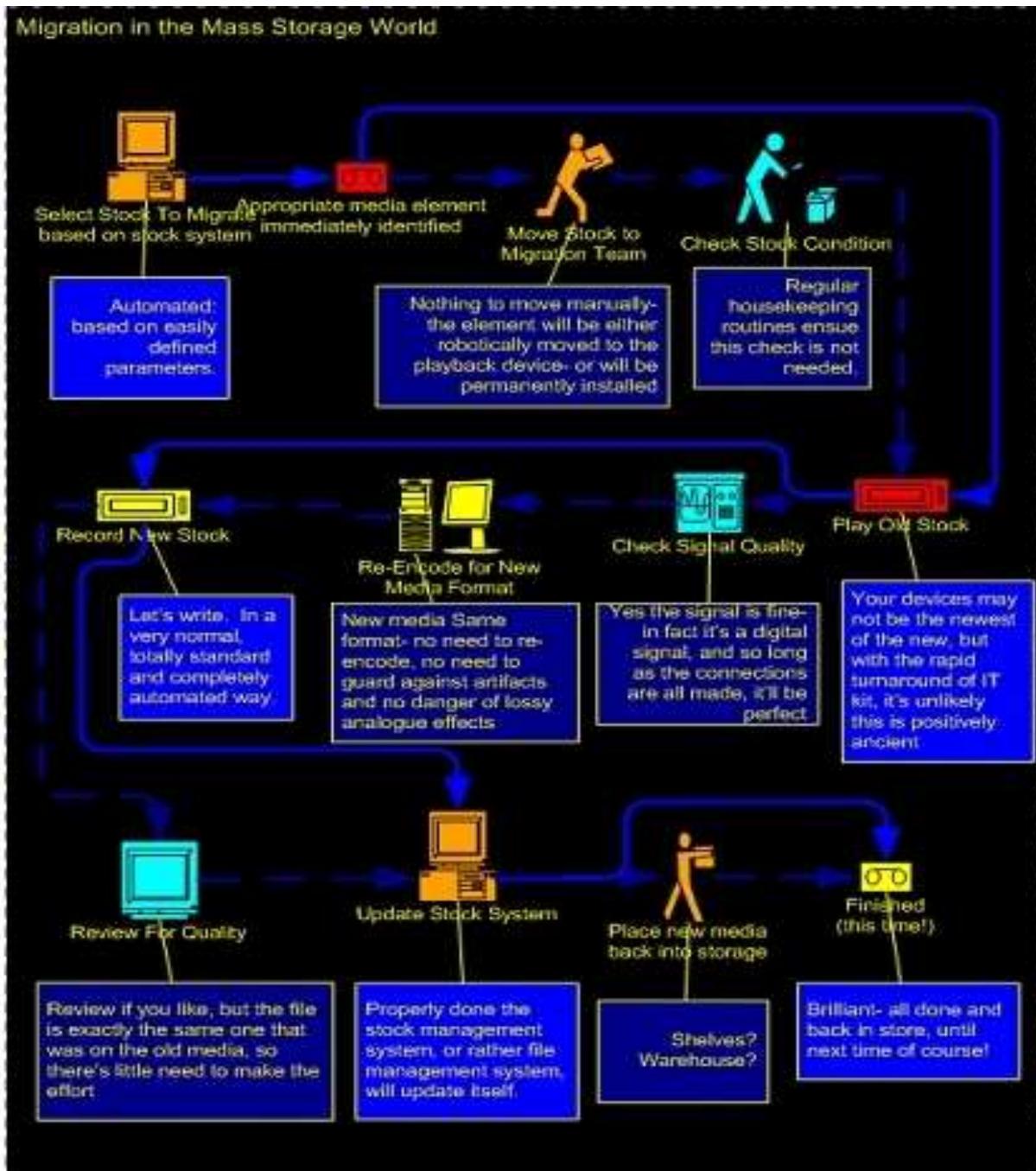
While there is much discussion of the exact details, the overall reduction in storage media cost has been remarkable. The cheapest terabyte external hard drive is now (Feb 2007) as little as US \$400, 1/3 the price 18 months ago¹⁶.

PriceHistory Feb-2007 (c) NexTag



Unfortunately, price per bit (or per gigabyte, to have a reasonable scale) should not be the primary consideration. If an archive had a choice of two buildings, one much cheaper but liable to be blown away or flooded or burned down – and the other in a very secure environment with an established safety history – it would be quite improper to make the choice strictly on cost. This is absolutely the situation with storage, and the advent of the terabyte external hard drive just means that the home user is now likely to lose a terabyte, whereas two years ago the loss would probably have been more like 200 GB. In other

words, five times cheaper means five times as much data on completely unsecured storage, which is not the direction in which archives should head. The next section gives information on how archives could and should cope with 'preserving the bits' – *digital preservation*.



5. Digital Preservation

5.1. Digitisation - solution or problem?

The main motivation for focussing this year's Audiovisual Status Report on *digital preservation* is because PrestoSpace is all about efficient and effective digitisation for audiovisual preservation, while the wider world now sees digital materials as a large, growing and unsolved problem. Further, recent studies ask whether digitisation (for books, photos and 'flat media') is in any way cost effective¹⁷.

The questions asked about the affordability and sustainability of digital storage for 'flat media' present two problems for audiovisual media:

- The amount of data required for audiovisual objects is much larger than for 'flat media', so any problems of affordability and sustainability will be magnified;
- Flat media can be microfilmed, a proven process with good technology and low cost. **There are no viable non-digital alternatives for audio and video material** – and the non-digital 'film as film' route for preservation of film collections is only viable so long as new blank film stock is being produced (which could be as little as another decade, with many types of film stock already out of production).

The issue of the cost of digital preservation will be taken up in Section 6.2 Cost of Digital Preservation, where evidence is offered supporting the view that digital preservation of audiovisual materials is affordable and sensible.

Digital preservation has been defined as: "*...the series of managed activities necessary to ensure continued access to digital materials for as long as necessary. Digital preservation...refers to all the actions required to maintain access to digital materials beyond the limits of media failure or technological change.*"¹⁸

This definition goes directly to the two main concerns:

- Maintaining access
- Media failure and technology change

PrestoSpace has raised exactly these issues in PrestoSpace Deliverable D13.4 Repositories Explained (see Reference 3), under the names *currency* and *persistence*.

- **Currency:** the ability to use repository content.
- **Persistence:** the ability to get content out of storage

5.2. Maintaining access

Maintaining access is about the currency or technical usability of content. Persistence is the ability to read back what was (once upon a time) written to a digital store, but persistence does not provide access – it is only a necessary condition. If the bits are in an obsolete file format, which cannot be opened and interpreted and played, then there can be no access.

Currency requires getting bits in a current format, meaning something that is suitable for the technical environment (computers and networks and operating systems and protocols and play out software) in use by the persons requiring the access.

There are two approaches to permanent access to content:

- 1) Migration: move the content from obsolete formats to current formats
- 2) Emulation: bring the mountain to Mohammed; put something on top of current computer technology (specifically, a Universal Virtual Computer) that supports exact emulation of antique technology, allowing antique data to be kept without change.

The interested reader can follow this debate in many papers¹⁹. There may be no single answer as to what is best. For data from databases, tied to specific applications, emulation may work and may be the only approach that could work. Emulation is a very challenging task, though one which is being attempted²⁰. As part of maintaining viability of old formats, both the US Library of Congress and the UK National Archive are making collections of software applications (such as word processors for a wide range of the computers and systems used in the 1980's and 1990's)²¹.

For audiovisual materials there should be, in general, no need to go to all the lengths required by emulation. Migration is a simpler process – *if the files themselves are simple*. To quote the Library of Congress (quoting me)²²: "Richard Wright at the BBC, has written, "PCM data, irrespective of sampling rate, word length, method of packing data into bytes and left-to-right or right-to-left arrangement of bits and bytes, can be decoded by relatively simple trial-and-error, and we can expect this to be the case indefinitely. PCM is in this sense a 'natural' representation for audio, and has very good long-term prospects regardless of the remaining problems of format migration." "

The point of the above quote is that simple files can be simply interpreted, so that *currency or maintaining access* is not difficult. The file can be interpreted fully, and if necessary re-written into whatever format is the current favourite for web browsers and other access technology.

PCM audio can be simple, but we are now moving into a range of more complex audio formats:

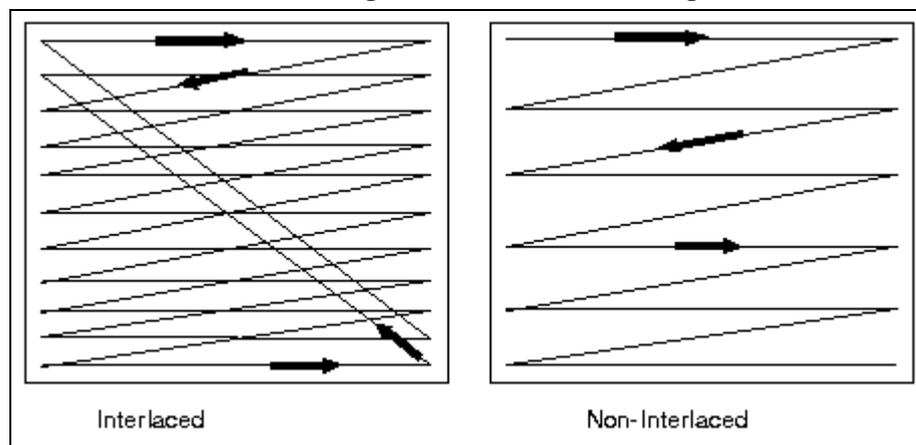
- encoded formats ranging from MP3 to Ogg (by way of WMA, APE, FLAC and others, plus and a variety of WAV files that, contrary to expectation, are not PCM but any of approximately 50 encodings registered with the WAV format)
- multichannel formats, generally D5.1 but that has its complexities, as it can be held in various fashions, including mixed in 24-bit stereo where 16 bits (per channel) are used for the two main signals, and the others are mixed into the remaining 8 bits. A more insidious complication is hard to imagine.
- formats with embedded rights management, almost exclusively proprietary – as an 'open DRM' would imply opening the code to manipulate the rights.

It is beyond the scope of this report, which is meant to be a status report rather than a technical brief, to go any further into the complexities just mentioned. However, the conclusion is clear. Rather than thinking of further complexity, such as a Universal Virtual Computer and registration of systems and software so that the above complex files could be played (if absolutely everything in all that complexity worked) – it would be far better for

archives to simply do everything in their power to keep such files out of their long-term repositories. This would only require another stage of migration – from non-standard to standard formats when data comes into the archive.

Digital video is inherently far more complex than digital audio. Colour video contains three simultaneous signals, plus audio (many professional video formats support 8 or more tracks of audio) plus timecode plus, perhaps, subtitling (closed caption information). The signal is not continuous as with audio, but is broken up (like film) into frames, each of which is usually subdivided further into interlaced fields, and each field is divided into lines. The allocation of bits to the three video signals is a parameter, which can vary (eg 4:2:0, 4:2:2, 4:4:4 and others). Other variables are: bits per sample, lines per field, frames per second and aspect ratio (which in turn affects the number of samples per line).

Figure 7 - Two types of raster scan, used to convert an image into a signal. Interlaced scanning allows twice as high a refresh rate, reducing flicker



All of that complexity is for unencoded signals, and there are very many ways to encode video, including again many which are proprietary. As an additional problem, most video is now originated as an encoded signal, such as some form of DV or some form of MPEG, and all video is digitally broadcast in encoded formats (MPEG again).

As with audio, the first issue is to reduce future complexity by converting to standard formats for video files within the archive. As encoding (compression) is guaranteed to cause problems in the future, there is a very strong case for not allowing compression. **PrestoSpace fully supports only using un-encoded, uncompressed video as the archive format for high-quality video signals.** For low quality, such as digital copies of VHS, PrestoSpace is working with the archives and with industry to select one 'mid quality' compressed format, which we can all agree upon and do our utmost to support and keep viable.

For digital preservation of video, archives need only create – from time to time – new access copies from their standard format, uncompressed master copies. This can be an entirely automatic process, and can be done 'on the fly' for many audio formats. Increases in computing power may soon make it possible to also encode video 'on the fly' – the whether or not that is the best approach is a complex system question.

To conclude this section, the advantage of the phrase “maintaining access” over the term “currency” is that it reminds us of the essential role of maintenance in digital preservation. Converting old file formats to new, current formats is one kind of essential maintenance (content maintenance), and must be in the budget and the management plan..

The other forms of maintenance refer not to the content, but to the storage and delivery system holding the content – and will be taken up next.

5.3. Media failure and system obsolescence

Media will wear out and fail, and become obsolete; audiovisual archivists are already very familiar with that fact. What may not be so well known, is that every single aspect of a digital archive has possibilities for failure, and every single aspect will definitely become obsolete – yet somehow the system as a whole is supposed to keep working. Here is the basic requirement for *persistence*: the bits have to endure, despite the fact that every component of the system holding the bits will have to be replaced.

The PrestoSpace emphasis on *persistence* may seem strange. Usually authors speak of ‘reliability of media’ or life-expectancy of storage and similar phrases. The definition given for *persistence* goes straight to the main issue: **reading back**. Anybody can write data to storage; the only interesting and difficult issue is whether the data can then be read back, and how accurately, and over what time scale.

The storage industry does not use the term *persistence*. Typically, storage industry information relevant to losing stored data is expressed in terms of error rates (of the data reading process), failure rates (at the device level) and media life expectancy.

There is a real gap here, because archivists have NO interest in read error rates and MTBF, and ‘digital archivists’ will also have no interest in media life expectancy. Meanwhile the storage industry provides data about storage media systems – the engineering data about the technology holding the data – and NOT information directly about the persistence of the content.

The storage industry provides the engineering data because it is known and is relevant to certain requirements, like system up-time and recovery time. When archivists talk to the storage industry, they should consider demanding the information they really want: will the content still be there in 20 years? Will it persist?

Specifically, archivists need to know:

- **How much content will be lost**, every year for N years. This is the one figure an archive can use to decide whether or not a storage strategy is acceptable.
- **The statistical distribution of the probability of loss**. This information allows an archive to assess the degree to which performance (of the storage strategy) can be trusted. It’s no good investing in a strategy with a 1% projected loss, if there is a 50% chance that the loss can be 10 times higher. This may look a bit complex and exotic – statistics about statistics – but it’s exactly the same complexity of information an insurance company uses to compute life assurance premiums. Only when an archive knows the confidence interval around the probability of loss can it make informed decisions about control of risk.

- **How the probabilities vary with N (how they vary over time).** This is again basic information, because storage strategies need to be re-assessed regularly. It may well make sense to change strategy after a shorter rather than a longer time, because probability of loss may well increase over time (or the costs – of keeping losses from rising – may themselves rise). A good horse to bet on can, in time, turn into a tired horse or an expensive horse. We are all familiar with this situation, especially with respect to being a car owner. Consumer guides to car ownership provide relevant information. Archives would like the same sort of information from storage providers.
- **How the probabilities vary with cost.** We all expect to 'get what we pay for'. We fully expect that a storage strategy with 99% persistence over 20 years would cost more than one with 95% persistence. How much more? Archives simply cannot get this information – not because the vendors won't say, but because the storage industry simply does not compute the statistics that the archivists most want to see.

The purpose of this section was to go into detail about the requirements of long-term archive storage using digital technology – and to show how much more archivists need to know about the pricing of storage, and what archives get for the price.

5.4. Where to learn more

Digital Preservation is now a recognised problem. The good news is that most of the issues apply to all 'bits in buckets', not just to audiovisual content. Therefore virtually all of the growing concern and literature is of potential help to audiovisual archivists. The bad news is that digital preservation is a problem, systems do fail, material will inevitably be lost – and it will probably take the pain of losses to prod us all into taking the full range of steps needed to prevent further loss.

The following are major online resources in digital preservation:

Digital Audiovisual Preservation:

AHDS 'Moving Pictures and Sound Archiving Study'²³

General Digital Preservation:

UK: Digital Preservation Coalition www.dpconline.org

Australia: PADI - Preserving Access to Digital Information www.nla.gov.au/padi/

USA: Digital Preservation (Library of Congress) www.digitalpreservation.gov/

Digital Preservation Tutorial www.library.cornell.edu/iris/tutorial/dpm/

Australia, France, the Netherlands, and the United Kingdom: National Digital Preservation Initiatives: An Overview www.clir.org/pubs/reports/pub116/contents.html

Europe: Digital Preservation Europe: <http://www.digitalpreservationeurope.eu/>

Germany: NESTOR Network of Expertise in Long-Term Storage of Digital Resources <http://www.langzeitarchivierung.de/index.php>

Canada: Canadian Association of Research Libraries

<http://www.carl-abrc.ca/projects/preservation/preservation-e.html>

6. Digitisation Costs in Europe

Cost is a complex issue, because the first question should always be not “how much does it cost?” but “what do you really want to do?” For proper accounting for all the costs to be included in a preservation project, the PrestoSpace cost model and online tools introduced in the previous chapters should be used.

However PrestoSpace has acquired data on digitisation, and those figures are given in the next section, 6.1. They are provided here in order to provide some comparison with the continuing costs of *digital preservation*, the particular form of *maintenance* that is required, forever afterward, once material has been digitised.

Costs of digital preservation are discussed in Section 6.2. There has been recent discussion about whether managed digital storage (an essential for digital preservation) is affordable, and hence whether digitisation is in any way advisable. As audio and video archives have no other choice (though film has), it is important to look carefully at digital storage and preservation costs. Section 6.2 presents a case for use and maintenance of digital technology that is quite affordable, in comparison to shelf-based storage.

6.1. Cost of Digitisation

VIDEO transfer cost data

Table 2 - PrestoSpace cost data for video transfers²⁴

PrestoSpace 2005	2004-2006	2007-2010
Estimated hours	211000	114000
Estimated costs (Euro)	27611500	18688500
Cost per hour (Euro)	131	164

Table 3 - Presto cost data for video transfers²⁵

Video Format Presto 2000	Average Cost per hour, €
2" all formats	176
1" all formats	357
3/4" U-Matic all formats	344
1/2" BETACAM and SP	230
1/2" MII	230
VHS	100

Table 4 - Example commercial transfer prices²⁶

SONY Preservation Factory, 2005
• 65€ per hr currently
• 50€ per hr projected
• 15€ per hr bulk discount (100,000+ Betacam SP)

FILM transfer cost data**Table 5- PrestoSpace cost data for film transfers (Ref 24)**

PrestoSpace 2005	2004-2006	2007-2010
Estimated hours	24900	11950
Estimated costs (Euro)	21315000	18962000
Cost per hour (Euro)	856	1587

Table 6 - Presto cost data for film transfers²⁷

Film Format Presto 2000	Average Cost per hour, €
35 mm negative	1994
16 mm negative	1685
16 mm reversal (all)	3011
16 mm reversal (tape joins)	2850
35 mm print	1900
16 mm print	1872
35 mm SEPMAG	2415
16 mm SEPMAG	373

AUDIO transfer cost data**Table 7 - PrestoSpace cost data for audio transfers (Ref 20)**

PrestoSpace 2005	2004-2006	2007-2010
Estimated hours	61000	(insufficient data)
Estimated costs (Euro)	8130000	
Cost per hour (Euro)	133	

Table 8- Presto cost data for audio transfers²⁸

Audio Format Presto 2000	Average Cost per hour, €
Shellac and vinyl, all formats	299
¼" tapes	238
DAT	195
CD	70
Weighted average	120

6.2. Cost of Digital Preservation

After digitisation, the digital data requires storage, and than storage must be managed (maintenance; digital preservation). This section considers the cost of maintained storage, looking at costs of a *service* rather than of a storage device or *product*. A product is essentially some hardware and some software, with a purchase price – but beyond that are the human costs around the storage system, with its staffing and maintenance that make the total system ‘trusted’, and give it any hope of permanence. We compare *managed storage* and with the raw costs of digital storage devices and the costs of shelf-based storage. This section concludes with a recommendation for the **implementation of repository functionality in an off-line (tapes on shelves) rather than online (everything on servers) environment**.

One of the first organisation to offer a ‘trusted digital repository’ service is the OCLC = Online Computer Library Center in the USA²⁹. The OCLC Digital Archive³⁰ is a commercial proposition, and uses the digital library standards (OAIS, METS³¹)

The OCLC service is an unlikely ‘maintained storage’ solution for digital audiovisual collections, because of the relatively high costs. According to the 2002 figures quoted by Chapman³², OCLC was charging US\$15.00/GB per year (at the best rate, for a terabyte or more of storage). Further, that charge is for secure digital storage, NOT for what Chapman refers to as “the capability to render intellectual content accurately, regardless of technology changes over time”³³.

The problem is: shelf storage had an annual cost in 2002 of about US\$10 per shelf foot (for storage with full environmental control). At 10 videotapes per foot, one hour each at Digibeta quality (80 Mb/s), the storage is 400 GB for \$4 (for the shelf) plus \$40 for the tape itself, meaning about \$0.10 per GB! The difference is a factor of 150. [This comparison is simplistic, because the tapes will last longer than one year, and so it is unfair to put the full \$40 Digibeta tape cost against the first year. But after five to ten years the tape will have to be migrated, and that cost has not been included in the figure of \$0.10 per GB. The two effects offset each other, though on balance the figure \$0.10 per GB is at the high end of estimates for shelf-based storage of videotape.]

Some further information on managed storage costs is given in the recent paper “The Digital Black Hole” by Jonas Palm of the Swedish National Archives³⁴. Managed storage requires staff as well as equipment, and staff can be a major part of the total cost. Palm quotes a Microsoft source on information management costs (primarily from the text world, where 1 TB of data represents roughly 250 million pages). In the financial world 1

TB of data requires one system manager, and other costs add up to US\$300k per terabyte. In other sectors this drops to one system manager per 10 TB, and for 'aggregators' like Google where data management is very automated, it drops again to one manager per 100 TB. Personal communication³⁵ with the Sun "Honeycomb" storage management project³⁶ also indicated a 'bottleneck at around 10 TB', with 'negative economies of scale' – meaning costs increased more rapidly as storage increased – for storage volumes greater than 10 TB.

These cost comparisons are summarised in the following table. The first row, 2002 costs, shows that managed storage is 150 times the price of storage of videotape on shelves, as discussed above (Chapman data). The middle column is for storage on hard drives of the cheapest sort – just the cost of the drives and nothing more.

Table 9- Estimates of annual audiovisual storage costs

Cost for 1 GB Year	Analogue on shelves	Digital media (offline)	Managed storage (online)
2002 (reported)	\$0.10	\$4	\$15 = 7 + 8
2006 (estimate)	\$0.11	\$1	\$11 = 7 + 4
2010 (estimate)	\$0.12	\$0.25	\$9 = 7 + 2
2020 (estimate)	\$0.15	\$0.02	\$7.05 = 7 + 0.06

As shown in the successive rows of the table, cost in the middle column drops dramatically. This drop is the storage variant of Moore's Law³⁷: storage capacity (for the same cost) doubling every 24 months. This drop has been steady for about 40 years. Prospects for the next 10 years are covered in another PrestoSpace publication: "Ten-year Forecast of Storage Evolution"³⁸

The managed storage figures quoted by Palm indicate that around US\$30k per terabyte, meaning US\$30 per gigabyte, was the "going rate" in 2005. These figures are higher than the Chapman figures, because they refer to IT industry in general rather than OCLC specific service for archive storage. But the difference between the Chapman and the Palm figures does not matter. What matters is:

- the figures are in rough agreement (\$30 and \$15 per gigabyte)
- the real issue is what happens to storage costs in the future.

Table 9 starts with the Chapman figure of \$15/GB/yr, and looks at what happens as Moore's Law applies to the raw storage part of that figure. We assume that about half the cost in the Chapman / Palm data represents staff and facilities³⁹, not storage devices themselves. These costs do not drop, but neither do they rise. Small improvements in IT management technology offset inflation, and basically IT systems cost now what they did 20 years ago – but the amount of storage in the systems (for that same price) has grown by the storage variant of Moore's Law⁴⁰: storage capacity (for the same cost) doubling every 24 months.

As seen in the right-hand side of the table, raw storage costs drop by Moore's Law, and all that is left, eventually, is the rest of the cost: management, maintenance, facilities. The table shows managed storage costs levelling off, not dropping as in the middle column. The numbers in the right-hand column could be off by a factor of two, but the general conclusion is that managed storage will NOT continue to reduce in cost. There will be little reduction after 2010, with cost stabilising at about \$10/GB/yr.

The first column, shelf costs, increases only as general inflation increases – which in most areas that are using digital repositories is quite low. The middle column follows Moore's law and shows storage becoming practically free – but only for raw storage devices, not managed servers.

The implications are very plain:

- archives will continue to find a managed service unaffordable, if by managed service one means a fully online system of mass storage (server-based storage; mass storage);
- letting materials sit as they are on shelves is untenable, as digitisation is required for preservation – certainly for audio and video materials;
- what's left is the middle column: use of raw storage media, which offers digital storage and low cost.

The problem with raw storage media is that it simply can't be trusted – for durations as short as three years much less the indefinite future. Raw media has NONE of the attributes of a trusted repository. Material just dumped onto storage media is very much at risk. The media is prone to failure, there is no inherent way to find anything, and the files themselves if they can be found and read (as data) will become obsolete and unplayable – and possibly unidentifiable and, effectively, lost.

The need, and really the only hope for the huge amounts of digital storage required by audiovisual archives, is to combine the physical storage approach of the middle column of **Table 9** – offline storage – with the functionality associated with the highest levels of service from online trusted digital repositories. Clearly what is needed are **trusted offline repositories**, with the security of the online repositories of the right-hand column of Table 1, at costs closer to the centre column.

PrestoSpace will support archives in convincing the storage industry that what archives need is to preserve content, not simply to write files to storage.

Cost of Digital Preservation

Table 8 gave the cost per hour of digitisation. Table 9 gives cost per year for various kinds of storage: tapes on shelves, datatape, and a full 'managed service' of high-end hard drives.

The following table gives cost per hour for digitisation and for ten years of storage, from 2010. While Table 9 had cost for managed online storage, taken from published data – the following table adds the cost for *managed offline storage*. The critical issue is how much offline storage can one 'offline manager' manage. As discussed, for online storage the number is somewhere in the region of 10 terabytes. For offline storage, we've assumed all the data has to be copied twice (in one decade). The copying requires equipment, and new media, and operators. The underlying assumption is that this five-yearly copying is the main work in turning raw storage into managed storage – plus keeping a full, separate copy in a separate location, and updating metadata when tapes are accessed and copied.

The estimate for "Managed storage (offline)" as compared to digitisation cost and "Managed storage (online)" is the whole point of Table 10, but there is little real data to go on. The formula used for estimating "Managed storage (offline)" is to add the 10-year digital media costs to an estimate of datatape duplication costs – and then multiply by four, for making two copies every five years. The estimate of datatape duplication costs is simply the videotape digitisation cost divided by 10 – because there is no evidence available to work at any greater precision than "order of magnitude" estimation.

Clearly datatape duplication costs could have almost no manual effort, but that would require expensive robotics. Either way, maintenance, the essence of digital preservation, has a cost – and at least Table 10 shows the implications of the stated assumption.

As is almost inevitably the case, the cheapest option is to leave analogue material on the shelf. We all know this. Unfortunately this option solves neither of the two main issues: preservation and access. Managed digital storage (offline) does address preservation, and use of web-quality online material addresses access.

Table 10 - Cost per decade for storage or digital preservation, compared with digitisation cost

Cost for 1 hr for ten years to 2020 -US \$	Digitisation⁴¹	Analogue storage, shelves	Digital media (offline)	Managed storage (online)	Managed storage (offline)
Video, full	164	15	37.8	4320	210
Video, DVD		7.5	5.66	648	23
Video, web		75	0.94	108	3.77
Film, full	1587	15	151	17280	660
Audio, full	133	7.5	1.32	151	65

If it were possible to maintain analogue audiovisual archives by copying old analogue materials to new, every 20 years, the estimate for 'analogue preservation' is given in the following table.

Table 11 - Estimate of 'analogue preservation' costs for a decade, compared with "managed storage (offline)"

Cost for 1 hr for ten years to 2020, US \$	Copying	Analogue preservation⁴²	Managed storage (offline)
Video, full	164	145	210
Film, full	1587	1141	660
Audio, full	133	108	65

Table 11 shows that digital preservation and analogue preservation cost about the same.

The major difference between the two approaches is that analogue preservation of audio and video is basically impossible, as there are no analogue formats that are not obsolete. It is welcome news to find that the above estimates place digital preservation back in the realm of reasonable costs.

A major flaw in Table 11 is with regard to the cost of analogue film preservation.

The table assumes a new analogue-to-analogue master copy is needed every 20 years, which is a defensible assumption for audiotape and videotape. However, a film master should easily last a century (if it isn't used, which is a bit of a problem). Table 11 could be made fairer to film, but there are complicated issues: how many prints to make from masters; whether to use 'interneg' as an intermediate between the master negative and the viewing prints. Those who really want to look into this detail should consult the excellent study by the Danish Film Institute covering all these issues, and outlining a 400-year, fully-costed preservation plan for film⁴³.

Archives and industry will need to focus on the *digital preservation* issue, and there should be a Competence Centre for defining the approach in detail, and supporting both industry and archives in the implementation. It is already difficult to introduce *trusted repository* concepts to the IT industry. For audiovisual material, these concepts have to also be introduced to the post-production and facility house industry – because that is where the files are produced. It will take a lot of support, from a Competence Centre because there is nothing else suitable, in order to lead facility houses through the into the realm of trusted repositories.

The need for *trusted offline storage* applies across digital libraries, but is especially urgent for audiovisual archives because of the size of audiovisual data. Other media can use online storage for their digitisation projects – though they then risk dropping into the black hole described in the Palm paper. Audiovisual archives cannot put full-quality video on mass storage in any significant volumes, and so need to move directly to secure offline storage. *Trusted offline storage* will then provide a lifeboat for rescuing other digitised media from the black hole effect – that online storage is unaffordable. Offline storage will hold large files of high data rate media, and online storage can be reserved for user access to low data rate proxies. This model has been in use for a long time, but without the essential factor of making the offline storage **trusted**.

An essential issue is identification of the content of a file, particularly with regard to understanding the file format, and ensuring formats remain viable. The important point, and a major motivation for a Competence Centre, is that format viability is not ensured locally. Identification is a local requirement, but once the file type has been labelled, the rest of the processes needed to ensure format viability can be implemented once, in one

place. These processes include maintaining players for the format, and establishing a migration path to new formats and new players. This aspect of digital preservation is not an item by item issue, but a general industry issue that can be managed in one centre of sufficient competence – a Competence Centre.

The conclusions of this section are:

- Digital preservation is affordable
- Offline digital media has a huge role in the future of audiovisual archives
- Management of such media needs to move from current practice to the processes established for digital repositories.
- There is a strong requirement for a Centre of Competence to guide archives and industry, and provide – centrally – the technology for a migration path for audiovisual offline content. This is the key technology needed for what Chapman referred to as “the capability to render intellectual content accurately, regardless of technology changes over time”⁴⁴.

7. Value of European Audiovisual Heritage

A cynic is a man who knows the price of everything but the value of nothing. Oscar Wilde

Lest the PrestoSpace project stand accused of cynicism, we present this chapter investigating the issue of value.

The phrase 'public value' is heard in discussions about access to media and cultural heritage content⁴⁵. The BBC has explicitly stated that one of its major functions is 'Creating Public Value'.⁴⁶ The BBC, in common with most organisations, defines Public Value in a general way, but does give a breakdown⁴⁷:

- *Democratic value* - to support active and informed citizenship by providing trusted, impartial, in-depth news and information that help people make sense of the world
- *Cultural and creative value* - to enrich the cultural life of the nation by enabling the UK's best creative talents to produce great original work and provide a broad range of memorable and enjoyable programmes and events
- *Educational value* - contribute to education for all by creating a wide range of accessible programmes and services that feed curiosity and enable people to learn throughout their lives
- *Social and community value* - help make the UK a more inclusive society by providing programmes and services that connect communities, encourage participation and help build a sense of place and belonging
- *Global value* - support the UK's global role by being the world's most trusted provider of international news and by showcasing the best of British culture to a global audience

There are three surprising facts about economists and assessment of Public Value⁴⁸:

- 1) It's what they do. Economists do not generally deal with commercial value – that is for accountants. One of the major functions of economic thinking is to deal with Public Value.
- 2) Economists measure public value in explicit monetary terms: Pounds Sterling, Euros, dollars and other currency.
- 3) Economists have been doing this since Adam Smith⁴⁹.

This chapter looks at value from an archivist's viewpoint, but quoting from leading legal and economic opinion.

7.1. Commercial and Public value

Archives need money in order to maintain their contents and continue to provide access – which is their function. The money comes from various sources, but one possibility is commercial sales. These sales have a *value* that is obvious and easily quantified. The problem with obvious things is that they can force less obvious issues into the background, even if the less obvious consideration is, really, the more significant.

Figure 8- A photo of a house – or a mountain?⁵⁰

There is a general idea that audiovisual archive content has to be protected from free access for two main reasons:

- There are underlying rights that have not been cleared
- Free access will destroy the possibility of commercial sales

In fact both these positions are assumptions, not reasons.

With regard to underlying rights, the assumption is that the rights-holders do not want public access. This assumption is often stated in virtually the same breath as a statement about the difficulty or impossibility of finding rights holders. Obviously the rights holders have not, in general, been consulted. The archive is making an assumption. However there are many examples of rights holders who would like nothing better than free access, or indeed any access, because a principal problem with many collections is that access is so very limited⁵¹.



Indeed one of the emerging phenomena is rights owners – creators of musical or literary works – who give their works away on MySpace⁵² or use other Internet-based mechanisms, in order to develop a following – or simply to be heard and read.

Examples are:

- “The Arctic Monkeys came to prominence through the MySpace site, and Lily Allen was signed to a records label after one of her songs proved popular on there.”⁵³
- Five collections of stories and novels by Cory Doctorow are available free from his website,⁵⁴ and Doctorow himself says *“enlisting my readers as evangelists for my work and giving them free ebooks to distribute sells more books. As Tim O’Reilly says, my problem isn’t piracy, it’s obscurity. Best of all, giving away ebooks gives me lots of key insights into how to make money without restricting the copying of bits. It’s a win-win situation”*.⁵⁵

With regard to protecting commercial value, the data from one major broadcast archive⁵⁶ is that, over the 70-plus year life of the archive, only 5% of the content has been exploited commercially. The consequence is that 95% of content is restricted, for the sake of the 5% that could gain commercial exploitation. This situation is definitely a case of seeing the house and not the mountain, in

Figure 8.

There is a much larger issue. The proper description of the situation just described is not “only 5% is of commercial value, so the rest may as well be opened for free access”. That would still assume commercial value is the only ‘real’ or at least economic and quantifiable value.

In the work of economists who do quantify public value, the situation with the archive just described is more like “only 5% has realised one form of monetary value, and that has prevented the rest from realising another form of monetary value which could be very much greater”. Under this second interpretation, it is more like “we sold the house, because we didn’t know we could make a national park out of the mountain, and spawn a range activities with a huge total monetary value.”

Pollock⁵⁷ gives the following examples: “The New York Times and The Guardian both give free access to their online edition but feature advertising as well as selling a corresponding print edition (complementary goods). ... And Yagan (2005) recounts his extremely successfully experience of ‘giving away’ content in the form of online textbooks first at SparkNotes and then, subsequent to its acquisition, at Barnes and Noble.”

Many more examples are in the Pollock paper. One of the most compelling is in the area of *public sector information* such as weather and geographical information. Such data are not generally seen as exciting when compared to audiovisual content, but the data are highly re-usable – if people are allowed access.

Quoting Pollock (Ref 57) on this data:

In the US, there is a strong tradition of placing public sector information in the public domain (Gellman 2004, Onsrud et al 1996). In Europe most countries restrict access to this data. For example, in the UK the records of parliamentary activity (Hansard) are under parliamentary copyright; government documents are under crown copyright; geographic data is owned by the Ordnance Survey, a trading fund, which prices on a cost-recovery (i.e. monopolistic) basis; weather data is controlled by the Met Office, which is another trading fund and therefore also prices at monopolistic rates. By examining the resulting outcomes for both commerce and society under these different approaches we can gain some measure of the benefits, or costs, of a public domain approach. The results are stark: a multitude of studies and research demonstrate the value of a public domain approach over a proprietary (cost-recovery) one.

For example, the EU’s own study, Commercial Exploitation of Europe’s Public Sector Information (Pira International et al 2000), found that although US governmental expenditure was only approximately twice as much as for the EU (9.5 vs. 19 billion Euros) the economic value generated was more than ten times as much (750 billion vs. 68 billion Euros). Though a rough and ready calculation, and allowing for the benefits of a unified market, this suggests a net value of a public domain approach to public sector information in the region of tens or hundreds of billions of euros – an enormous sum.

The conclusion from these examples is that realising *public value* should become the major consideration for audiovisual collections. The next section looks at methods for releasing that value.

7.2. Releasing value

It is out of place for a sub-chapter of a report on audiovisual preservation status to explain the economic workings of the modern world in any detail, so what follows is no doubt simplistic – but there is much more information in the references.

Only two types of value will be considered: *commercial* and *public*.

Commercial value is based on selling goods and services. The issue with regard to archives is the changing nature of the market. Anthony Lilley refers to the conventional market as the “age of scarcity” – with media industries concentrating on a limited range of high-turnover, big market products. Lilley describes “the skills of the age of scarcity – hitmaking, hype and cross-media promotion among others.” [Ref 53, p7].

This is a marketplace of restriction and protection:

- restricted range of products that have to ‘earn their shelf space’
- restricted access: come to our theatre or cinema, buy our recording or publication, tune in to our channel – because we are ‘sole source’
- restricted rights: look but don’t touch, passive viewing and listening but no active involvement. Difficult to ‘quote’ media even for academic purposes that are allowed under copyright legislation; illegal and very difficult to make copies or extracts for general purposes.

The new marketplace created by computers and the Internet is altogether different. This market has no shelves, it is as easy to stock 100,000 items as to stock 100 items, and the major issue is ‘presence’ (standing, trust, brand recognition). This is a market requiring the “skills of the age of plenty, the exploitation of the so-called long-tail, the importance of authenticity of voice, and the power of communities of interest.” [Ref 53, p7]

In these new markets, the goods which previously couldn’t get on the shelves are transformed into major sources of profit. The whole ‘long tail’ phenomenon is a recognition that there is more money to be made for ‘worst sellers’ than from ‘best sellers’. The ‘worst sellers’ are at the low frequency-of-sales long tail of the distribution of items (according to sales rate). Amazon the online bookstore (and now books-and-everything-store) is often sited as making more money from the long tail than from the popular goods⁵⁸, and EBay has achieved its position (profits of US\$310m on turnover of US\$1.4bn for one quarter – Q2 – in 2006⁵⁹) by offering a way to sell the longest tail of all: all of our used goods – a market that had been considered insignificant.

These markets work by attracting people to ‘use’ the site, a new combination of goods and services. It could be compared with the “come to our national park” approach to commercial sales, where people do come to use the park, but they also buy things.

The situation is complex, because the most successful new company of all, Google, doesn’t sell goods or services. It provides a free service, which attracts users. The users come to Google as an intermediary point in their quest for information (in the widest

sense) in this new 'activity place' called the Internet. Google then sells pinpoint advertising space, tuned to the wants of the users. The traditional corner shop had an element of this model, when it charged 20p for people to put cards in their front window. But every visitor to the shop had to search all the cards to see if any were of interest. The Google Corner Shop waits for a customer to ask the shopkeeper for a magazine about cars – and then all the cards selling cars or car repairs come flying out of the shop window and hover about the customer.

The point for audiovisual collections is: people like audiovisual content. If audiovisual collections were online, people would come to the Audiovisual Archive Shop, to look around – and if enough people come then that creates the kind of opportunity that Google has so effectively exploited. The point about the long tail is that in an online world there is no such thing as a 'worst seller'. The obscure or specialist collection can have its day.

For the obscure archive to 'have its day' via the Internet, it must be:

- online; this requires digitisation, encoding and hosting
- public – because any restriction on access is a proportionate restriction of interest, visitors to the site – and ultimately any value, public or commercial, associated with the site
- visible – the hardest condition of all. As in the Cory Doctorow quote [Ref 55] "my problem isn't piracy, it's obscurity". Ways to become visible are a huge subject, but the above two points, online and public, are preconditions.

Public value is released in a very simple way – the material is simply opened to public access. The commercial market requires skills of merchandising and mechanisms of brand building, including a great reliance on advertising. Public value is much more straightforward: it merely requires removal of restrictions.

Commercial and public value differ diametrically in their approach to access. For commercial markets, access is restricted – to release commercial value. Restriction may be entirely necessary for manufactured goods and 'commodities'; for the new information market it can become a lose-lose-lose situation.

- The economy gets much less benefit than if public value were also realised;
- The general public get restricted access;
- Rights owners do not enter the new markets (those based on internet access and the 'long tail'), and generally lose out as the old markets dry up.

There is a particular danger for heritage material coming, originally, from the commercial sector. This was dramatically illustrated by the fate of cinema films in the last century:

"For the first few decades of the 20th Century there were few official audiovisual collections. Film stayed with the companies that produced it. Many such companies lasted only a few years. Their accumulated negatives were seen as 'used goods' of small value, and so roughly 2/3 of the film production pre-1940 no longer exists"⁶⁰

The example shows the disastrous consequences of allowing commercial value to dominate public value: short-term commercial value is at odds with long-term heritage considerations. The 'restrict, then dump' commercial drivers end up opposed to both preservation and access.

Where commercial value depends upon restricted access, public value requires enlarged access, which can then produce a win-win-win situation.

- Economic benefit is maximised, certainly for media which is part of the ‘information industry’, where the public value dominates (if allowed to!);
- The general public get enlarged access;
- With the enlarged access, rights owners can, if they wish, develop new business models that can be far more remunerative – in commercial value – than is the case for the old models.

These are strong claims, so the next section provides some examples.

7.3. Interactions

The following are examples of increasing public value, and increasing commercial value at the same time.

1) **Giving it away** – and selling more: the Prelinger archive was one of the first to provide web access. This is an archive of ‘non-fiction film’, a wide range of the sorts of films that you see anywhere except in a cinema: educational, industrial, government, public information and all other forms of non-cinema film, which Rick Prelinger collected to preserve, and because he was interested. He had a small commercial business, but the collection was not well known or well used. In his own words:

We’ve a historical stock footage collection which began in 1983 as a hobby and two years later turned into a business and today has grown to become a very large collection containing 51,000 titles and 30,000 cans of unedited footage.

Here’s our old model of scarcity: strict, physical control of our material. Nothing went out unless it had timecode on it, unless somebody paid a licence fee. We were very rigid about asking people to report. Tons of VHS cassettes and sometimes films would come to us and we would painstakingly count frames and in the case of compositing or any kind of layering or image manipulation, we would total up with great glee how many of our shots appeared at the same time in a frame so that we could double/triple bill.

...

So at the beginning of 2001 we made a page with links to 270 films which has now grown so that now [2005] there are about 2000 films available online for free viewing, downloading and reuse.

Here’s the characteristics of what we’ve put on-line: all the films are downloadable, shareable and re-usable... So what’s “free” is a download. What you pay for is a physical copy which is often higher res. You get a Creative Commons Licence at no cost with the download. If you want a specific written agreement with your name on the top which you’re going to need if you’re doing any kind of production that requires an Errors and Omissions Policy, then you have to pay;

In 1999 there were 200 stock sales ... in 2004 there were 400. We saw very, very little unpaid production usage in 1999. Our estimate for 2004 is 800 uses in production. The ten public screening events in 1999 were all done by me but in 2004 only six of about 70 events were. But in 2004 the number of films distributed to the public has risen to 800,000 and that's really the dramatic statistic, a statistic that we're extremely proud of. Again, this does not count material that's FTP'd or delivered through BitTorrent or other means so actually the total is higher. The total number of our films distributed online is just about 3 million.⁶¹

In summary, when Prelinger put free content online, his commercial business doubled, and the non-commercial business went from almost nothing to 800 productions and 800,000 downloads (viewings).

2) **Trusting the public:** as well as giving it away (in reduced quality), the Magnatune⁶² music label lets the customers buy full quality albums, and pay anything they like (from US\$5 upwards). They recommend that users pay US\$8. Surprisingly, the average payment is US\$8.20. The free downloads enhance public value, and this almost anarchic sales strategy has increased commercial value by 0.20 per album. The company has been growing since its launch in 2003, and the full 2006 yearly statistics are available online⁶³.

3) Turning to public value on a national scale, and with specific regard to audiovisual heritage, the Dutch national audiovisual archive B&G = Beeld & Geluid (or NISV = National Institute of Sound and Vision, a PrestoSpace partner and steering board member) has worked with the Dutch government on the value of national audiovisual content to the nation's 'information market'. The result has been a major award – the largest every, to anyone, anywhere – for audiovisual preservation. The funding is €174 million, and Dutch government economists have put together a public value business case anticipating a net benefit to the nation of around €60m. This is a major breakthrough: the nation's heritage preserved, and making money out of it!

The B&G is to be congratulated by all archivists, all governments and by anyone who has every listened to radio, watched television or even seen a photograph – because they have shown leadership to all of us on how to really address preservation of the audiovisual heritage. Further information is in the reference⁶⁴.

Any country wishing to follow the lead of the Dutch, and any nation which thinks its audiovisual heritage is worth as much as that of The Netherlands – can use the following table as a guide to what the government should fund⁶⁵

Table 12 – National Audiovisual Preservation Budgets: How Much to Ask For

Country	Bid, € millions
Austria	88
Denmark	58
France	606
Germany	805
Ireland	55
Italy	545
Netherlands	173
Portugal	70
Spain	361
Sweden	93
U.K.	642

7.4. A call to arms

'He who receives an idea from me, receives instruction himself without lessening mine; as he who lights his taper at mine, receives light without darkening me'
(Thomas Jefferson 1813)

There is a special and perverse situation about many audiovisual archives: they restrict public value for the sake of little or commercial value. Archives can be remarkable backward about even trying to extend access.

Common phrases used in such archives are:

- “prevented by copyright”;
- “prevented by cost of clearing copyright”;
- “prevented by fear of consequences of failure to clear copyright”;
- and prevented by – *something* – from even testing those consequences.

What prevents archives from even trying to put material online, or to show or distribute ‘orphan’ material with no clear copyright and ownership? Curatorial conservatism? Fear of the copyright holders? What about fear of the failure to serve the public, failure to realise public value, and failure to prevent loss of unique and irreplaceable material?

The argument of Rufus Pollock indicates that the general public value (user value) being lost is very likely to be far greater than the specific commercial value gained – at that is without including loss of the material itself, when lack of access causes lack of funding and hence lack of preservation. The “sin of omission” to test new forms of access has clear consequences: *that which could be saved will be lost*, and with it all hope of realising any value, or any sort, by any model or measure.

Why press for public access to audiovisual archives? Because it adds value and saves archive content. Restricted access loses value, and loses archive content. The issue is not one of breaking the law, it is one of breaking out of unnecessary and outmoded habits.

There is a caricature of an archivist or librarian as someone who wishes the public wouldn't come in and disturb and damage the collection. In reality, it is the role of the archives is to protect content, but the worst way to protect it is to deny access – because that denial generates no value, and leaves archives unable to afford to run preservation projects. *Protecting the content requires public interest, which only comes from public access.* Copyright holders need to be challenged, educated and wherever possible recruited to the advantages, all around, of maximising access.

This section of chapters on cost and value concludes with a quote from *Le Monde Diplomatique*, from an article printed to coincide with the campaign⁶⁶ led by the Institut National de l'Audiovisuel (leaders of PrestoSpace) to call for world attention to the problem of endangered audiovisual media:

*The history of the audiovisual heritage of the 20th century is largely based on chance: certain things happened to survive, and we're lucky to have them. If this heritage is to be maintained through the 21st century and into the future, the attitude must shift from chance to intent. Thorough planning is required, and **cultural heritage institutions must include preservation funding within their standing budgets.***

*In return, **these collections and institutions must provide far better access.** There is no technical reason why all this material cannot be available in every citizen's home, via the web. The rights and legislative barriers are already under pressure, and will have to come down – to free this material and provide uniform, democratic access.*⁶⁷

8. Access

The following is a short summary of audiovisual access initiatives in 2006.

8.1. Public Access and User-Contributed Content

The biggest news of the last year was the explosion of interest in a single website, which 'simply' provided a place for people to deposit, for public access, audio and video materials: YouTube.

There have been heated discussions about what YouTube is, whether it's good or bad, and where it might lead. But the basic facts are clear: YouTube is a phenomenon:

From August 2006⁶⁸:

- In a single month the number of videos on the site grew 20% to 6.1 million
- YouTube has some 45 terabytes of videos
- Video views reached 1.73 billion

Additional statistics reported in July 2006⁶⁹ were:

- 60% of all videos watched online are on YouTube
- 100 million downloads a day
- 65,000 uploads a day (that's 1,538 downloads per upload)
- 20 million unique users per month
- 29% of US online multimedia entertainment market

And of course the biggest statistic of all was the sale of YouTube, which was only 18 months old and was losing huge amounts of money (to pay for the download bandwidth) – to Google in October 2006 – for US \$1.65 billion⁷⁰.

If YouTube remains 'the place to be seen' for audiovisual materials, the archives should seriously consider using this free, available and popular website as a front window for their own holdings – not necessarily for hosting all their online content, but certainly for providing "tasters, teasers and trailers" which can then lead people to the archive's home page and full offering. If archives do pursue the YouTube option, the major task, as already mentioned in other contexts, is to get noticed – discovered – in all the other material. It is well beyond the scope of this survey to offer guidance on 'how to raise your internet profile' – but there are many sources of information and guidance⁷¹.

8.2. Archive Access

Another success story took place firmly in the world of audiovisual archives: the Institut National de l'Audiovisuel (INA) launched *Archiv pour tous*⁷² in April 2006 and it was an immediate hit.

Reported usage went to 5 million visits a day, up by roughly a factor of 100. The service is free streamed viewing, and a small charge for downloads:

“About 80 percent of the collection is free. For copyrighted material, INA charges €1 to €3, or \$1.29 to \$3.86, for the purchase of a 48-hour viewing window or €1 to €12 for full downloads.”⁷³

In last year’s “Audiovisual Preservation Status”⁷⁴ report there was a list of other archives offering (or planning to offer) large amounts of their collection free to the public for online viewing. The entry point to the category of “major online audiovisual archive” appears to be 10,000 or more hours, and INA is firmly in that category, with plans to grow by 5,000 hours per year.

The BBC has announced similar plans, with the mid-2007 launch offering being around 1000 hours, but again with ambitious plans to get to 30,000 hours in three years.

The Netherlands launched a range of new services in 2007, with the opening of the new building of the B&G in Hilversum. They have an online portal to their collections,⁷⁵ which runs on their new catalogue IMMIX which has been several years in development – and they have an online introduction to their *Experience*⁷⁶, which introduces their achievement:

“Nowhere else in the world is a public facility like the media experience directly linked to the archives”.

It’s also quite a nice building:



B&G are providing public access to 10,000 hours of digitised media as part of the Experience. With this initiative, B&G remind us all about the significance not just of public access via the web, but of the continuing importance of ‘people through the door’.

These major European initiatives show that:

- audiovisual archives are coming into real prominence;
- opening large amounts (thousands of hours) of archive material for public access is possible;
- these actions attract huge amounts of interest.

8.3. Commercial Access

This report has stressed public value, but that does not exclude consideration of commercial value. We have discussed the problem of perspective, where a narrow commercial interest obscures a potentially much larger public interest. But providing that a full perspective can be maintained, it is not PrestoSpace's role to in any way discount or discourage commercial activities.

An important study on commercial possibilities of small audiovisual archives was completed during 2006, and is available online⁷⁷.

Five commercial case studies are reported in detail, and the conclusions of the report identify six "key critical success factors for commercial exploitation of audiovisual collections":

- *Cataloguing*
- *Technical viability*
- *Rights*
- *Good knowledge of all aspects of the care and management of audiovisual media*
- *Investment*
- *Markets and Marketing*

Full details are in the reference just cited.

8.4. Research

There is a range of research on audiovisual access, but some key projects have just started, and a new call for further research has also just started.

DPE- Digital Preservation Europe⁷⁸. This project started in April 2006. It follows from ERPANET, and says it "*fosters collaboration and synergies between many existing national initiatives across the European Research Area. DPE addresses the need to improve coordination, cooperation and consistency in current activities to secure effective preservation of digital materials*".

MEMORIES⁷⁹ This three-year project started June 2006. It concerns using audio for semantic indexing of archive content. It has an impressive set of partners, including the technical specialists Memnon, a number of archives – and UNESCO.

SEMEDIA⁸⁰ This is a STREP, meaning a relatively short and small project, which started January 2007. From the website "*The overall objective of SEMEDIA is to create new methods, environments and widely usable tools for media labeling, searching and retrieval from very large collections ...*"

Finally, the next major round (*Framework 7 = FP7*) of European funding has just started, with the first call for proposals having a submission date of 8 May. There is funding for audiovisual preservation research, and a range of digital library activities. Further information from the EC IST website⁸¹ or from national support bodies.

9. European and Global Support for Audiovisual Preservation

9.1. EC support for audiovisual preservation

In the context of the i2010 Digital Libraries initiative, the European Commission adopted (24 August 2006) a Recommendation on digitisation, online accessibility and digital preservation of cultural resources.

The Recommendation calls on EU Member States to set up large-scale digitisation facilities, so as to accelerate the process of getting Europe's cultural heritage online via the European digital library. It also calls for action in various other areas, ranging from copyright questions to the systematic preservation of digital content in order to ensure long term access to the material.

The full text of the Recommendation and other relevant information are available from the website of the i2010 Digital Libraries initiative⁸².

9.2. TEL and i2010 Digital Libraries and EDL

The EC support of libraries, digital libraries and audiovisual preservation under the TEL (The European Library) and i2010 Digital Libraries was covered in last year's report.

Two significant developments since then are:

- Minerva – the effort at the European level to coordinate library and archive digitisation actions – has been extended. Starting from October 2006 MINERVA Project is enlarged to MINERVA EC⁸³, Ministerial NETwork for Valorising Activities in digitisation, eContentplus.
- EDL – the European Digital Library⁸⁴. This is a long-awaited development. There are many institutional digital libraries (some institutions have more than one!) and a few national digital libraries, but the open issue would be how to bring these together across Europe. It is a very important issue, because YouTube has shown that what makes an online collection important is the level of its usage, not the quality or uniqueness or cultural significance of its content.

The EDL started in September 2006, so it is really quite new. It is a project, not an institution. PrestoSpace is also 'just a project', so we know full well the insecurities of that status, especially when working in an area where we are striving for long-term effects.

EDL states its position as working "towards the integration of the bibliographic catalogues and digital collections of the National Libraries of Belgium, Greece, Iceland, Ireland, Liechtenstein, Luxembourg, Norway, Spain and Sweden, into The European Library".

The integration of the bibliographic data (the catalogues) of these members is already the task of TEL, The European Library. So the new business of EDL must be the "integration

of the ... digital collections". This is a laudable and necessary task, and PrestoSpace has offered to work with EDL to provide support in the area of audiovisual materials (and in storage and distribution).

However important European libraries are significantly absent from the EDL membership, and one might well wonder what they are expected to achieve over 18 months, on an EC contribution of €1 million. A far greater level of support will be needed if Europe is to truly collect and present its digital library materials in a manner that will compete with commercial or advertising-driven offerings from outside Europe.

9.3. Google and Euro-Google

Mention of 'advertising-driven offerings from outside Europe' that are involved in audiovisual and digital libraries must call to mind Google, with their Google Books initiative, Google Maps, Google Video and now You Tube as well.

Last year's Preservation Status report mentioned Quaero, sometimes described as Europe's answer to Google (or at least, Europe's response). According to Wikipedia:

*"Quaero was announced by Jacques Chirac during the French-German ministerial conference of Reims in April 2005, and was scheduled to be officially launched in early 2006 by the Agence de l'innovation industrielle (All). Close to €90m from the governments of France and Germany are supposed to go towards development of Quaero"*⁸⁵

Recent reports⁸⁶ describe the projects as 'scrapped' owing to insufficient investment to create a meaningful alternative. But the Wikipedia article [Ref 85] goes on the mention "a semantic search project called Theseus that will be distinct from Quaero"⁸⁷ and is evidently largely a German initiative.

9.4. The European Archive

One promised development that did launch, on 28 September 2006, was The European Archive⁸⁸. They describe themselves as follows:

The European Archive is a digital library of cultural artifacts in digital form. We provide free access to researchers, historians, scholars, and the general public.

They are related to the Internet Archive in the USA (and Alexandria, Egypt), and share a common dedication to providing access – essentially to 'anything that belongs in a library'.

A full description of the launch and of the goals and methods of The European Archive is on their information pages⁸⁹.

They are already working with other European cultural heritage and archive institutions, notably – yet again – Beeld and Geluid. They list 22 'movies' [British Government Public Information Films], 321 recordings [Public Classical Music Beeld en Geluid] and 4 web collections [Italian Domain Snapshot 2006] on their main page – a good start.

They want to “become an essential piece in the European cultural heritage landscape”, so any audiovisual archivists who want to make contact with world-leading access technology should consider working with the European Archive.

9.5. The Global Dimension

UNESCO World Day for Audiovisual Heritage

UNESCO General Conference has resolved that 27 October each year be proclaimed the World Day for Audiovisual Heritage. A feasibility study on the introduction of the Day was conducted in conjunction with intergovernmental and non-governmental organisations active in this field.⁹⁰

UNESCO's own report on the first **World Day for Audiovisual Heritage** is on their website⁹¹.

Further News from UNESCO (from the IASA email list):

The Convention on the Protection and Promotion of the Diversity of Cultural Expressions, adopted by the General Conference at its 33rd session 2005 will enter into force on March 18, 2007. After the very difficult negotiations to set up the text of the Convention its ratification by more than 30 member states in a very short time is a great success.

Audiovisual Preservation Training Workshops

International audiovisual organisations continued to organised conferences and training courses in preservation during 2006.

FIAT/IFTA held conferences in Rabat⁹², Morocco during June, and in Phnom Penh⁹³, Cambodia during December.

IASA held its main conference in September 2006 in Mexico, where European experts could exchange information with a whole range of delegates from North and South and Central America⁹⁴.

10. Service Industry Developments

10.1. Vienna Workshop

PrestoSpace held a major workshop in May 2006 in Vienna, devoted to matching the needs of audiovisual archives to the services (and prices) available from industry. This was the main PrestoSpace contact with the services industry during 2006, though a range of informal contacts continue. There was a large attendance (about 70 delegates) for such a specialist meeting.

One major issue discussed was: what happens to the relationship between archives and industry after PrestoSpace? Daniel Teruggi of INA outlined a proposal around the idea of a Centre of Competence which would have an independent, long-term existence (possibly within one or more host institutions), and which would provide a range of services to keep archives and industry working together – for higher quality, lower prices and more preservation (better, cheaper, faster).

A full report is online⁹⁵. The day was deemed so useful that PrestoSpace will do it again (same city, same month, different building), again with a FIAT meeting just preceding it. Further information is below.

The Competence Centre is not a PrestoSpace invention. The Netherlands already operates a form of that model, with Beeld and Geluid providing the audiovisual service. Project Minerva, which of course is now Minerva-EC, has a list of de facto competence centres⁹⁶ – those this seems to be a list of institutions which *have* competence, rather than a list of places to go for a systematic service.

Since Vienna there has been much discussion, including at the European level concerning Competence Centres for all the processes relevant to digital libraries and archives. Further information will be given at this year's Vienna User Meeting.

10.2. Storage Service Provider offerings:

When audiovisual materials are digitised, the result has to be stored somewhere. PrestoSpace has the SAM (Storage and Archive Management) part of the project specifically to provide information in this area. It is a difficult area for many archivists, because providing massive amounts of digital storage within an archive is a whole new technology and discipline. One alternative is to let somebody else do the storage, as a managed service.

The interesting proposition is for the same company that “made the bits”, the facility house or digitisation specialist, also store the bits – and create a website and an online catalogue and run the whole thing – if required.

This full-service model has been taken up by several of the companies who have taken an interest in the PrestoSpace project. It's not the role of PrestoSpace to single out

companies for favouritism – but this report wouldn't make much sense if we didn't name any companies at all.

A list of the ones that have talked to SAM – which should not be taken in any way as an endorsement or as a full list, are listed below (alphabetically):

Moving Media ⁹⁷	Ireland and UK	in association with Cambridge Imaging ⁹⁸ and TVP ⁹⁹ . Digitisation, storage, delivery
Memnon ¹⁰⁰	Belgium, UK, France	Also offering a full digitisation, storage, delivery service
Ovation Data Services ¹⁰¹	UK and USA	Data tape and storage specialists to the audiovisual and petrochemical industries
Stream UK ¹⁰²	UK	Mainly hosting for web delivery

PrestoSpace would like to hear from other companies offering managed storage services.

MAY 11-15th 2007 Vienna: FIAT Seminar and PrestoSpace Service Workshop
--

Meeting Notice: We don't have a meeting notice! FIAT will be meeting on 11-12 May and PrestoSpace on 14-15 May, with Sunday free for sightseeing. Both the FIAT and the PrestoSpace meetings will be in the same venue, in central Vienna – not at the ORF TV building further out.

[By the time this report is released (early April, after review) this notice should be updated.]

11. References

- ¹ Presto D2 Archive Preservation and Exploitation Requirements, Table 4.1.1 page 9
<http://presto.joanneum.ac.at/Public/D2.pdf>
- ² Presto D3.2 Key Links Systems Specification, 2001. Figure 1, (generic preservation) Reference Process ; p3. The PDF is online, http://presto.joanneum.ac.at/Public/D3_2.pdf
- ³ What Archives Want – the requirements for digital technology. www.ebu.ch/en/technical/trev/trev_308-archives.pdf;
also PrestoSpace report D13.4 Repositories Explained (awaiting review before public release)
- ⁴ Preservation Guide:
<http://wiki.prestospace.org/pmwiki.php?n=Main.WhatIsPreservation#Maintenance>
<http://wiki.prestospace.org/pmwiki.php?n=Main.PreservationStrategy#Conservation>
- ⁵ From the PrestoSpace online tutorial for "migration transfers" <http://prestospace-sam.ssl.co.uk/>
- ⁶ Deliverable 2.1 User Requirements Final Report; Preservation and Digitisation Plans: Overview and Analysis. 2004
http://www.prestospace.org/project/deliverables/D2-1_User_Requirements_Final_Report.pdf
- ⁷ <http://www.dpconline.org/docs/events/050726psxpiti.pdf> and
<http://www.dpconline.org/docs/events/050726psxpbbc.pdf>
- ⁸ <http://www.prestospace.org/sam>
- ⁹ PrestoSpace – KDCS – TAPE - BUFVC Workshop: Audiovisual Preservation for Culture, Heritage and Academic Collections, London Nov 2006.
<http://wiki.prestospace.org/pmwiki.php?n=Main.TAPE-London>
Preservation Cost models – Matthew Addis, Southampton University
<http://wiki.prestospace.org/uploads/Main/Addis.pdf>
- ¹⁰ http://www.prestospace.org/project/deliverables/D5-3_Difficult_Media.pdf
- ¹¹ <http://wiki.prestospace.org/>
- ¹² Magnetic Tape Storage and Handling A Guide for Libraries and Archives, Dr. John W.C. Van Bogart National Media Laboratory, 1995 http://www.clir.org/PUBS/reports/pub54/estimation_of_LEs.html
- ¹³ 80% goes through OK, the remaining 20% takes four times longer which is equivalent to another 80%, yielding a total project time and cost of 160% (as compared to a no-problems cost of 100%).
- ¹⁴ Figure taken from "Planning the digitisation, storage and access of large scale audiovisual archives" Matthew Addis, Freddy Choi, Ant Miller. 2005 <http://eprints.ecs.soton.ac.uk/12231/>
- ¹⁵ http://en.wikipedia.org/wiki/Moore's_law
- ¹⁶ <http://www.nextag.com/LaCie-Big-Disk-Extreme-77082608/prices-html?nxtg=2d4c0a1c0520-4AFE0D9954AD3CA8>
Price history : <http://www.nextag.com/LaCie-Big-Disk-Extreme-77082608/price-history-html>
- ¹⁷ "The Digital Black Hole" Jonas Palm, Swedish National Archives, 2006
www.tape-online.net/docs/Palm_Black_Hole.pdf
- ¹⁸ "Counting the Costs of Digital Preservation: Is Repository Storage Affordable?" Stephen Chapman Weissman Preservation Center, Harvard University Library, Cambridge, MA. Journal of Digital Information, Volume 4 Issue 2, 2003. jodi.tamu.edu/Articles/v04/i02/Chapman/
- ¹⁹ Jones, M and Beagrie, N, 'Preservation Management of Digital Materials: a Handbook', British Library, London, 2002
- ²⁰ Emulation for JPEG photos : http://www.kb.nl/hrd/dd/dd_onderzoek/uvc_voor_images-en.html
Digital Preservation Strategies: http://en.wikipedia.org/wiki/Digital_preservation
- ²¹ 'Emulation: Context and Current Status'
www.digitaleduurzaamheid.nl/bibliotheek/docs/White_paper_emulation_UK.pdf
- ²² PRONOM: nationalarchives.gov.uk/pronom/ ; LOC Digital Formats: digitalpreservation.gov/formats/
- ²³ Audio and Video Preservation Reformatting - A Library of Congress Perspective
<http://www.archives.gov/preservation/conferences/papers-2003/fleischauer.html>
- ²⁴ Report will be available from JISC, but a copy can be obtained from richard.wright@bbc.co.uk
- ²⁵ From data in PrestoSpace Deliverable 2.1 User Requirements Final Report. P30.
- ²⁶ Presto D2, 2001 Archive Preservation and Exploitation Requirements; Table 4.1.3 Average costs for preservation of video formats; <http://presto.joanneum.ac.at/Public/D2.pdf>
- ²⁷ Announced at FIAT/IFTA conference, New York, Autumn 2005

²⁷ Presto D2, 2001 Archive Preservation and Exploitation Requirements; Table 4.1.2 Average costs for preservation of film formats; <http://presto.joanneum.ac.at/Public/D2.pdf>

²⁸ Presto D2, 2001 Archive Preservation and Exploitation Requirements; Table 4.1.4 Average costs for preservation of audio formats; <http://presto.joanneum.ac.at/Public/D2.pdf>

²⁹ www.oclc.org

³⁰ www.digitalarchive.oclc.org/

³¹ Digital Repositories and related standards are fully covered in PrestoSpace D13.4 Digital Repositories Explained, submitted for project and reviewer approval in January 2007.

³² jodi.tamu.edu/Articles/v04/i02/Chapman/

³³ jodi.tamu.edu/Articles/v04/i02/Chapman/ p4

³⁴ www.tape-online.net/docs/Palm_Black_Hole.pdf

³⁵ Personal conversation with Mike Davis, **Honeycomb** senior project manager. Sun Honeycomb have participated in PrestoSpace STAG (Storage Technology Advisory Group) events

³⁶ research.sun.com/minds/2005-0628/

³⁷ en.wikipedia.org/wiki/Moore's_law

³⁸ www.prestospace.org/project/deliverables/D12-5.pdf

³⁹ This cost breakdown is deduced from the "one staff member per ten terabytes" statement quoted by Chapman. If storage costs around \$100k per \$300k for 10 TB, and includes one full-time staff member, then it is reasonable to conclude that about half the cost is for the member of staff and for all associated infrastructure.

⁴⁰ en.wikipedia.org/wiki/Moore's_law

⁴¹ Using the PrestoSpace estimates for transfer costs per hour or material: film and video use 2007-2010 estimates, audio uses the 2004-2006 data because of insufficient data for 2007-2010

⁴² cost of storing two copies on shelves, plus half the cost of a transfer (because a 'refresh' transfer is needed every 20 years) -- except using 0.7 instead of 0.5 times the cost of a transfer, to allow for inflation. So

⁴³ Nissen, Larsen, Christensen & Johnsen, Eds, "Preserve then Show", Danish Film Institute, 2002, ISBN 87-87195-55-0).

⁴⁴ jodi.tamu.edu/Articles/v04/i02/Chapman/ p4

⁴⁵ http://www.ksg.harvard.edu/news/opeds/2004/crabtree_revolution_started_in_library_new_statesman_0927_04.htm

⁴⁶ BBC, June 2004, Building Public Value - BBC report on renewal of its charter
http://www.bbc.co.uk/thefuture/pdfs/bbc_bpvp.pdf

⁴⁷ www.bbccharterreview.org.uk/pdf_documents/Attitudes_to_BBC_radio_and_public_value.pdf

⁴⁸ Rufus Pollock, private conversation. Also economists prefer the term *social value*.

⁴⁹ en.wikipedia.org/wiki/Adam_Smith

⁵⁰ Stodertal Valley, Flickr, Creative Commons licence; <http://www.flickr.com/photos/swq/398764643/>

⁵¹ It would be remiss of PrestoSpace to pick on individual archives, but the author has visited several national collections which average less than ten visitors per day.

⁵² www.myspace.com

⁵³ Inside the Creative Industries: Copyright on the ground. Anthony Lilley IPPR July 2006
www.ippr.org/publicationsandreports/publication.asp?id=483

⁵⁴ www.craphound.com/novels.php

⁵⁵ <http://www.craphound.com/bio.php>

⁵⁶ Name temporarily withheld pending permission from their Press and Publicity office

⁵⁷ The Value of the Public Domain, Rufus Pollock IPPR July 2006
www.ippr.org/publicationsandreports/publication.asp?id=482

⁵⁸ The Long Tail, Chris Anderson. Wired 12.10, Oct 2004 www.wired.com/wired/archive/12.10/tail.html

⁵⁹ www.theregister.co.uk/2006/07/20/ebay_Q2_2006/

⁶⁰ Film Preservation Study, National Film Preservation Board, 1993.
<http://lcweb.loc.gov/film/filmpres.html>

⁶¹ http://creativearchive.bbc.co.uk/news/archives/2005/05/transcript_rick_1.html

⁶² <http://magnatune.com/>

⁶³ http://blogs.magnatune.com/buckman/2006/12/end_of_year_sta.html

⁶⁴ http://www.futureofthebook.org/blog/archives/2006/11/dutch_fund_audiovisual_heritag.html
palimpsest.stanford.edu/byform/mailling-lists/amia-l/2006/10/msg00054.html

⁶⁵ in proportion to relative size of national economies (GDP; public value is in some sense tied to the information market of the country, not the population or the size of the archive or any 'cultural' metric).

⁶⁶ International Appeal for the Preservation of the World Audiovisual Heritage
<http://www.fiatifta.org/aboutfiat/policy/petition/index.php>

67 Richard Wright: Le Monde Diplomatique: Oct 2004 **The Audiovisual Heritage of the 20th Century – the Accidental Archive.** [**Online in French : "Le dilemme des archives"**
<http://www.monde-diplomatique.fr/2004/10/WRIGHT/11571>]

68 http://www.micropersuasion.com/2006/08/youtube_by_the_.html

69 http://open.typepad.com/open/2006/07/0_60_in_under_1.html

70 <http://www.washingtonpost.com/wp-dyn/content/article/2006/10/09/AR2006100900546.html>

71 for instance: www.imediaconnection.com/content/10847.asp (and thousands more)

72 www.ina.fr/archivespourtout

73 <http://www.ptvdigitalarchive.org/2006/05/25/keeping-up-with-the-bbc-archives-pour-tous/>

74 <http://www.prestospace.org/project/deliverables/D22-6.pdf>

75 <http://portal.beeldengeluid.nl/>

76 <http://experience.beeldengeluid.nl/>

77 <http://www.mlasoutheast.org.uk/archives/southeastarchivessector/avcommercialstudy/>

78 <http://www.digitalpreservationeurope.eu/>

79 www.memories-project.eu/

80 <http://cordis.europa.eu/ist/kct/semmedia-intro.htm>

81 www.cordis.europa.eu/ist/about/fp7.htm

82 http://europa.eu.int/information_society/activities/digital_libraries/index_en.htm

83 <http://www.minervaeurope.org/>

84 <http://edlproject.eu/>

85 <http://en.wikipedia.org/wiki/Quaero>

86 *The Economist*, January 6th 2007 (pp. 5); quoted in the Wikipedia article on Quaero

87 IT-Gipfel: Quaero heißt jetzt Theseus; <http://www.heise.de/newsticker/meldung/82708/from/rss09>

88 <http://www.europarchive.org/>

89 <http://www.europarchive.org/launch-official.php?PHPSESSID=0ae87a13d979f28629161942d1ff7514>

90 <http://www.ccaaa.org/newsarticle.php?id=49>

91 http://portal.unesco.org/ci/en/ev.php-URL_ID=23341&URL_DO=DO_TOPIC&URL_SECTION=201.html

92 http://www.fiatifta.org/conferences/seminars/past/rabat_2006/index.html

93 http://www.fiatifta.org/conferences/seminars/past/cambodia2007/cambodia_report.html

94 <http://www.iasa-web.org/iasa0009.htm>

95 http://www.prestospace.org/user_group/20060518/Memo-PS_Preservation_Factory_Workshop_18-05-06.pdf

96 www.minervaeurope.org/interoperability/competentcenters.htm

97 www.movingmedia.tv/

98 www.cambridgeimaging.co.uk/

99 www.tvparhive.co.uk/

100 www.memnon.be/

101 www.ovationdata.com/

102 www.streamuk.com/