KEEPING
AUDIOVISUAL
CONTENTS
ALIVE

Deliverable 2.1.1
Audiovisual preservation strategies, data models and value-chains

SEVENTH FRAMEWORK
PROGRAMME

Matthew Addis (IT Innovation Centre)
Richard Wright (BBC)

Document administrative table

Document Identifier	PP_WP2_D2.1.1_preservationstrategies Release 1.01
Filename	PP_WP2_D2.1.1_preservationstrategies_R0_v1.01.doc
Workpackage and Task(s)	WP2 Models and environments for long-term audiovisual content preservation WP2T1 Models for audiovisual preservation
Authors (company)	Matthew Addis (IT Innovation), Richard Wright (BBC)
Contributors (company)	Christoph Bauer (ORF), Johan Oomen (B&G), Marcel Mattheijer (B&G), Laurent Boch (RAI), Marius Snyders (B&G).
Internal Reviewers (company)	Beth Delaney (B&G) Ido Peled (ExLibris)
Date	04/03/2010
Status	Delivered
Type	Deliverable
Deliverable Nature	R - Report
Dissemination Level	PU - Public
Planned Deliv. Date	31 December 2009
Actual Deliv. Date	11 February 2010

Abstract	This is a report on preservation strategies, models and value-chains for digital file-based audiovisual content. The report includes: (a) current and emerging value-chains and business-models for audiovisual preservation; (b) a comparison of preservation strategies for audiovisual content including their strengths and weaknesses, and (c) a review of current preservation metadata models, and requirements for extension to support audiovisual files.

DOCUMENT HISTORY

Version	Date	Reason of change	Status	Distribution
0.01		First Draft	Outline	Confidential
0.06	26/11/2009	Added section on preservation of video through migration	Draft	Confidential
0.08	27/11/2009	Added section on multivalent	Draft	Confidential
0.11	4/12/2009	Added case studies for business models and value chains (BBC, ORF), added descriptions of audio and video formats (BBC)	Draft	Confidential
0.14	10/12/2009	Added cost models	Draft	Confidential
0.15	12/12/2009	Finished off Sections on business models and preservation strategies	Draft	Confidential
0.17	13/12/2009	Finished Section 3 – draft now complete	Draft	Confidential
0.18	14/12/2009	Finished executive summary and fixed formatting problems	Draft	Confidential
0.19	30/12/2009	Revision of B&G case studies	Draft	Confidential
0.21	9/2/2010	Final minor corrections	Final	Restricted
1.00	11/2/2010	Final release	Delivered	Public
1.01	25/2/2010	Update to include NDIIPP report	Delivered	Public

Table of contents

Scope

The European Commission supported PrestoPRIME project (www.prestoprime.org) is researching and developing practical solutions for the long-term preservation of digital media objects, programmes and collections, and finding ways to increase access by integrating the media archives with European on-line digital libraries in a digital preservation framework. This result will be a range of tools and services, delivered through a networked Competence Centre.

This is a report on preservation strategies, data models and value-chains for digital file-based audiovisual content. The report includes:

(a) a review of the current and emerging value-chains and business-models for audiovisual preservation;

(b) a comparison of the preservation strategies for audiovisual content including their strengths and weaknesses, and

(c) a review of current preservation metadata models, and the requirements for their extension in order to better support audiovisual content.

The objective of this report is to provide guidance and quantitative comparison on the various approaches that can be used to *keep audiovisual files safe and accessible over long periods of time*. A long-time[1] in this context means over a long enough period for changes in technology, formats, or user needs to be a cause for concern. Safe[2] means an acceptable level of content loss over time, and accessible means the ability for the designated community[3] (content users) to be able to retrieve and understand the content.

There is no 'one size fits all' preservation strategy for audiovisual material. Different content owners, users and custodians have different requirements for the cost, safety, longevity and accessibility of audiovisual material. Given the variety of different approaches that can be taken, the issue becomes one of which option to choose for a given set of circumstances. Currently, this decision is often based on ad-hoc assessments, 'gut feel' and sometimes conflicting recommendations by various experts in the field.

This report aims to take make the identification and selection of a preservation strategy for audiovisual content a rational and methodical process supported by quantitative cost models that take some of the guesswork and uncertainty out of what to do.

With the help of this report, we anticipate that both archives and their service providers will be in a better position to assess, select, monitor and review their preservation strategies. This includes consideration of the wider issues of economic sustainability and whether to implement digital preservation of AV files as an in-house activity, in collaboration with other archives, or by outsourcing to third-parties.

---

[1] For example, see OAIS for a definition of 'long-term'
http://public.ccsds.org/publications/archive/650x0b1.pdf
[2] Arguably no loss is acceptable, but from a practical standpoint, cost-effective preservation isn't perfect. For example, Richard Wright at the BBC proposes a realistic 'safety' target of less than 1% loss over 20 years.
[3] Again see OAIS definitions.

Executive summary

This report covers three main aspects of preservation strategy:

1. The value chains and business models that underpin preservation strategy in the widest sense, i.e. addressing economics and sustainability. This addresses the need to understand where the money will come from, how to keep it coming, and how preservation and access should be done as a result.

2. Specific technical preservation strategies, e.g. migration and multivalent, for keeping files alive and accessible over the long-term. The focus here is on file formats (encoding and wrappers) and the long-term total cost of preservation, especially storage. These are areas where archives are bewildered with choices and have difficulty making objective and justified decisions on what to do.

3. The metadata that needs to be captured to enable long-term preservation. A valid preservation strategy has to capture at the outset all the information required to be able to take the correct actions to preserve content. But what is this metadata and how should it be structured?

The specific questions that this report addresses include:

- What file formats should I use? What storage technologies make sense?

- What is the total cost of preservation of AV files and how is this broken down?

- What preservation metadata do I need to capture and keep?

- What examples are there of how preservation can be sustained by monetising the archive or providing public access?

- What business models for preservation as a service are emerging and are archives considering these?

- What are the challenges and costs of preservation at scale?

- What metadata standards should I consider and where do they fall short?

- Are there different strategies I can use for different types of content?

To answer to some of these questions, the authors reviewed the practices and plans of several national archives and broadcasters within the PrestoPrime consortium along with the existing literature to reach the following main conclusions:

Preservation, Access and Sustainability are all interconnected and a strategy for one can't be developed without considering the others.

The strategies for preserving content (e.g. which formats to use, or whether to do the work in house of use a service provider) need to be considered in the context of

who needs to access and use the content along with the economics and business models of long-term sustainability.

Sustainable preservation through online commercial sales or public access is not an easy business model to make work in practice

There are many people who have tried and failed to use content licensing as a way to generate revenue to sustain preservation. More successful are those who combine commercial licensing with public access services. This typically aligns with an archive's core mission of providing access to its content whilst at the same time recognising that monetisation can bring in complementary and additional revenue streams.

Digital preservation and archive hosting services are in their infancy, but they are starting to emerge. This area will grow quickly over the next few years.

Third-party services for long-term hosting and access to digital AV assets are currently very few and far between. However, for many small to medium archives who are ill equipped to cope in a world of digital files, the outsourcing of preservation to third-parties is an attractive, although currently unproven model. Beyond the issues of trust (see D3.4.1), the economics of remote data centres (power, cooling, economies of scale, network accessibility) along with the need for immediate online access to archive content are likely to come together to provide strong drivers for more widespread uptake of digital archives as services.

The issues of preserving audiovisual content in a given file-format are complex, although the overall long-term strategy is simple: uncompress.

Selecting a suitable preservation strategy requires consideration of the issues of sustainability, in particular tool support, the ways in which digital content can be stored and the Total Cost of Ownership over time of this storage, and how the content needs to be accessed and used. Our simple recommendations are to adopt an open wrapper format (e.g. MXF), use uncompressed audio bitstreams (e.g. BWF or AES3), then the following for video:

- If the video content is already in digital file format and is sustainable then you should keep it unchanged.

- If you have to change format and can afford uncompressed storage, then this is the first choice.

- If you have to change format and can't afford uncompressed, then lossless compression using an open format, e.g. JPEG2000 is the next best thing.

- If you have to change format and can't afford lossless compression, then choose a lossy format that is long-lived[4] and open with a quality level that matches the source material.

---

[4] Long lived means long enough for storage costs to fall to the point where uncompressed formats are viable. This will typically be less than 10 years.

The total cost of storage for digital files is far higher than the cost of the media used and has to consider trends as well as regular migration.

Current options for cost-effective long-term storage at any scale nearly all require automation through mass storage devices (disk servers and tape robots). The short obsolescence times for the components of these systems, including media, require continual migration and active management. Traditional models of 'media on shelves' are unlikely to be cost effective except for extreme cases where the need to access the material is very infrequent, e.g. once a year or less. Our simple set of rules for calculating long-term storage costs are:

1.   Take today's cost for commodity storage media (disk drive or data tape) and multiply by 10 to get the total annual cost of managed storage for that media.

2.   Repeat step 1 for multiple copies, e.g. one copy on disk and one on tape.

3.   Multiply the result of step two by 4 to give the TCO for a lifetime of storage.

The cost of providing Internet access to audiovisual content can be many times higher than the long-term cost of storing that content.

For archive content that is regularly accessed, e.g. where there are more accesses per month than items of content, then the cost of providing Web-based access, including all the necessary infrastructure (e.g. transcoding, hosting, networking), even for 'browse' quality content, can be higher than the cost of preservation, including master copies. This has major consequences in terms of investment, business models and sustainability.

Compressed (encoded) audio or video data formats are idiosyncratic by definition, which needs to be taken into account during preservation.

A file management or preservation process that needs knowledge of the structure of data within an encoded file will need the same information as in an application that plays or renders such a file. Given the large number of encodings for audiovisual data, plus the rate at which new ones develop and old ones become obsolete, it is an open question whether effort should be put into identifying a separate, formal data structure for any encoded audiovisual data. Instead, the effort could go into keeping the bitstream (binary representation; data sequence) intact, and keeping a valid and working player (codec) for each such file type.

Technical metadata can fully describe audiovisual content, so that there is no dependence (in principle) on players and environments. This approach is perfectly simple for unencoded audio, relatively simple for unencoded video, and probably not at all possible for encoded signals.

The audiovisual industry would benefit from much more agreement on where and how to place metadata in proprietary file types, and on ensuring that *all* the decode parameters were part of that metadata.

1. Value-chains and business-models for AV preservation

In this section we cover two main areas:

- Value-chains and business models that **get the money** to do preservation and provide access, for example commercial sale of content as a revenue stream or public access services supported by government funding.

- Value-chains and business models used to **do the preservation**, for example outsourcing to service providers for storage, transcoding and access. Here we focus on new and emerging models rather than common practice.

These two areas are of course not independent. For example, for a small archive the main business driver might be online access to generate the funding/income to sustain its preservation work. In this case, it could make sense to use remote digital preservation and hosting services for the content. This makes online access simpler and avoids the need for in-house skills and infrastructure, not just for preservation but also content access and distribution.

1.1.   *Overview and sources of further information*

There is plenty of literature about the range of business models and strategies that could be applied to preservation and access for cultural and commercial content (images, video) etc. Much of this information has emerged as a result of attempts by many to sustain online access to multimedia content. For example, in 2003 The Multimedia World Watch, International Federation of Multimedia Associations (FIAM) published a report on "Ensuring the sustainability of online cultural and heritage content: from an economic model to an adapted strategy"[5] This report marks an early and comprehensive effort to pull together the issues of value, business models and risks, all of which have a major part to play in developing a strategy for sustainability.

More recently, 2009 has seen two key publications in this area, firstly a report on "Sustaining Digital Resources: An On-the-Ground View of Projects Today. Ithaka Case Studies in Sustainability"[6] by Ithaka and more recently the Blue Ribbon Task Force (2009) interim report on Sustainable Digital Preservation and Access (BRTF-SPDA)[7]. Both explore the issues of sustainability in terms of economics (especially the costs), value (including how to create it) and business models (where the money will come from). Whilst the discussion in both reports is extensive and provides help to those who face similar challenges for sustainable audiovisual preservation and access, it is the inclusion of case studies, with real numbers, that is perhaps the real value of these reports.

The Preserving Public Television Project[8] in the United States, which is supported by the Library of Congress through the National Digital Information Infrastructure Preservation Programme (NDIIP)[9] has just released a report on Strategies for Sustainable Preservation

---

[5] 'http://www.fiam.org/en/FIAM-M2W-Study%20online%20content.pdf'
[6] http://www.ithaka.org/ithaka-s-r/strategy/ithaka-case-studies-in-sustainability/report/SCA_Ithaka_SustainingDigitalResources_Report.pdf
[7] http://brtf.sdsc.edu/biblio/BRTF_Interim_Report.pdf
[8] http://www.thirteen.org/ptvdigitalarchive/
[9] http://www.digitalpreservation.gov/

of Born Digital Public Television[10]. This excellent report is notable in that it provides a very up-to-date and comprehensive look at sustainability both in terms of where the money comes from (e.g. funding sources and business models) as well as the detailed costs of establishing and running a preservation infrastructure.   An insight into the content of the report is shown in Figure 1.   Therefore, PrestoPRIME strongly recommends the reading of the PPTP report alongside this report, the Ithaka report and the Blue Ribbon interim report.

Community Model
Rights holders pool resources to financially sustain a repository to provide preservation services

Submitter Fee-for-Service
Rights holders pay a repository to provide preservation services on a fee-for-service basis

Corporate Sponsorship/ Advertising
Commercial entity funds the repository directly, or indirectly supports preservation through ads

Public Sponsorship/ Philanthropy
Public body or philanthropist funds the repository to provide services that further the public good

User Fee-for-Service
Users pay the rights holder(s) and/or the repository to access preserved digital content at the point of use

Endowment
An endowment is created to fund long-term operations

Joint Venture Between Rights Holder and Archive
Rights holder(s) and repository enter into a financial agreement in exchange for ongoing access and/or preservation services

Services/Consulting
Services and technical assistance are provided to outside organizations on a straight contract basis to produce revenue to support preservation
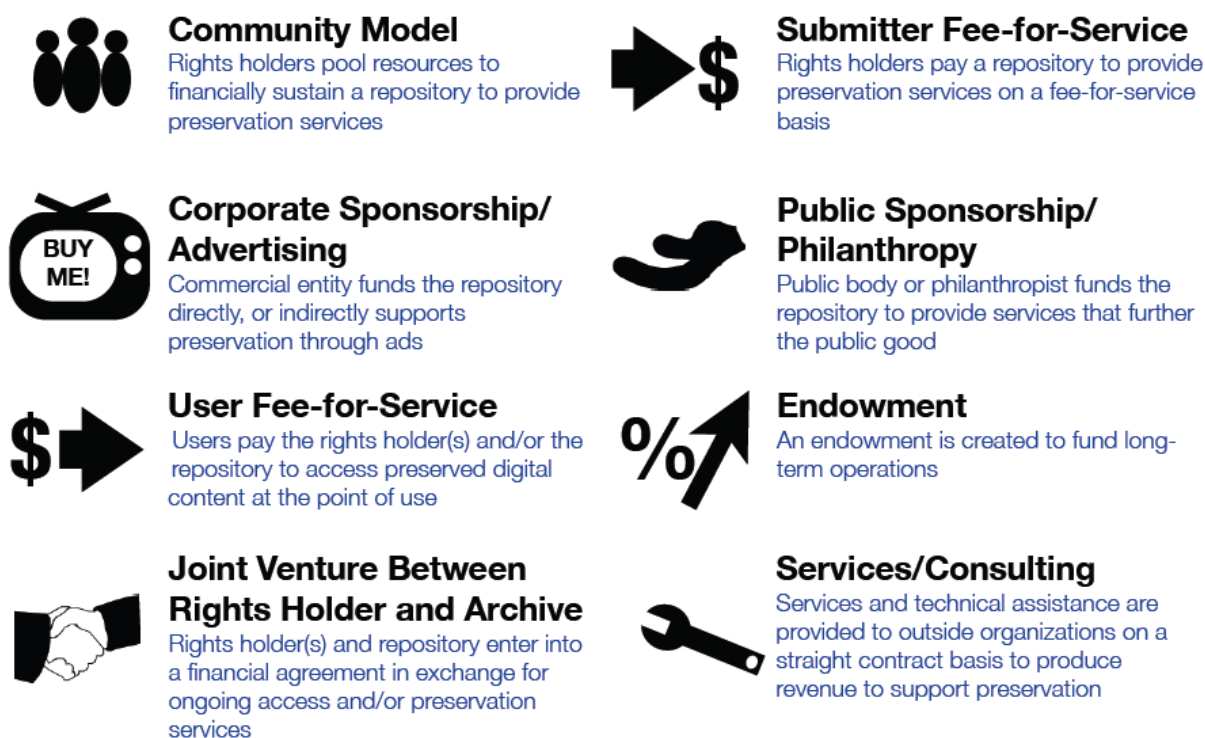
Figure 1 revenue models discussed in the Preserving Public Television Project report on Strategies for Sustainable Preservation of Born Digital Public Television[11].   The figure above is copyright of the members of the project has been reproduced directly from the report.

The Ithaka, Blue Ribbon and Preserving Digital Public Television reports can also be used in the context of more general work on the issues of sustainability, value and securing funding. Here the work of the eSPIDA[12] project is helpful for two reasons. Firstly, it presents a lifecycle for going from strategic objectives of an organisation through to project proposals, investment decisions, and on to delivery and performance evaluation. This helps frame the process that many archives need to go through in securing and using investment into preservation. Secondly, eSPIDA describes a balanced scorecard approach to value (based on Kaplan and Norton's model[13]), which allows non-monetary aspects of value (customer satisfaction, brand awareness, higher quality content, faster access, new services, flexibility and others) to be considered in a comparable way to the traditional financial aspects of cost savings, licensing, further investment etc. This is particularly relevant to many audiovisual archives where many aspects of value are intangible and

---

[10] http://www.digitalpreservation.gov/
[11] http://www.digitalpreservation.gov/
[12] http://www.gla.ac.uk/espida/index.shtml
[13] http://www.balancedscorecard.org/BSCResources/AbouttheBalancedScorecard/tabid/55/Default.aspx

hard to quantify in monetary terms (e.g. as described in detail in the UK DEMOS[14] report on cultural value[15]).

Rather than further enumerate the already excellent work available on sustainability, we conclude this section with a set of specific case studies of what has worked, and what hasn't, in the experience of the audiovisual content holders in PrestoPrime. It is specific examples such as these that the community is crying out for as beacons on the way forward and as contact points for further information.

We also note that 'content rights' (moral, intellectual, copyright etc.) and 'rights clearance' and 'rights management' operations can be a major bottleneck in the exploitation of archived audiovisual content, whether analogue or digital. In many cases the original contracts need to be analysed and interpreted by specialists to determine if reuse is allowed, which considerably increases costs and causes delays: vast portions of the archive remain unused due to uncertainty over rights. Rights can apply to different parts of the content, but as contracts with rights holders are often drafted before the content is produced, there is often no explicit association between the resulting shots and rights. Furthermore, rights can have a temporal span and are often reacquired or renegotiated subsequently. Different regulations apply across Europe, expressed with different terminologies, with different approaches based on different legal traditions.

A recent study by RAI has identified as many as 800 different types of rights associated with its assets. The sheer number of rights (which will scale at the European level) illustrates how difficult it will be to create an automatic rights system, and how necessary it will be for the provision of public access to archive content and the ability to run a business on it. While a lot of literature exists on the definition of Digital Rights Management (DRM) systems to convey and enforce copyright statements in the B2C marketplace, where the licensing models are comparatively simple, there is little research into models that could cover the range of rights found in the real world in a professional environment.

Instead of attempting any further analysis of this complex area, we refer the reader to work done by PrestoPrime specifically in the area of rights (one of the tasks of work package 4) that will become publicly available in the second year of the project along. There are, and will be further, discussions and presentations on this issue from the PrestoPrime partners, e.g. at the various public workshops that the project is hosting.

---

[14] http://www.demos.co.uk/
[15] http://www.demos.co.uk/publications/culturalvalue

1.1.1.   Case study: Putting the BBC archive online

Drivers:

The main interest is public value. This has been a consideration ever since the original plans of Greg Dyke in 2003[16] which resulted in the Creative Archive. The focus was a "service … free and available to everyone, as long as they were not intending to use the material for commercial purposes."

The issue is releasing public value, in material which has been paid for by the licence-fee payers, but "Up until now this huge resource has remained locked up, inaccessible to the public because there hasn't been an effective mechanism for distribution."

The whole issue of public value was presented, in a broadcast archive context, in the PrestoSpace annual audiovisual preservation progress report for 2007 - Deliverable D22.8 Annual Report on Preservation Issues for European Audiovisual Collections (2007) – pp27-35.

"Ultimately, this archive could be the biggest and richest service the BBC produces." : Mark Thompson, Director General Prospect Magazine 4th July 2009[17].

Funding:

Funding online public access is a case of striking the appropriate balance between the public service and commercial exploitation objectives of the BBC. This is against a backdrop where no industry-wide business model has yet emerged.

The BBC's current strategy is thus to build on the existing activities allowed under the Trust Service Licences. This avoids committing large amounts of licence fee funding whilst the reach, quality, impact, value and market context become clearer. BBC Worldwide are also continuing to explore commercial exploitation of the archive where reasonably possible.

The increased public service uses of archive content are likely to contain strongly themed propositions based around landmark strands or complementing online activities such as the NHU Wildlife Finder launched in September 2009. Similar 'clip-based' propositions are being developed and cleared for launch in 2010 as part of the Knowledge Strategy, including a Science Portal and an Arts Finder.

The BBC has allocated a fixed budget to the release of publicly funded content, set at a maximum level of 0.5% of each Group's content budget per year, and this will fund any rights clearance costs.

Sustainability:

Essentially archive access becomes a new BBC "offering" – whose sustainability is backed by the BBC itself. The general sustainability of BBC TV, radio and Internet offerings is

---

[16] http://news.bbc.co.uk/1/hi/3177479.stm
[17] http://www.prospectmagazine.co.uk/2009/07/markthompsononthefutureofthebbc/

backed by BBC policy, agreed up to board level. Programmes and Internet activities do come and go, but the public service remit – and the level of public support as evidenced by the Creative Archive and Open Archive trials – mean that online archive access should be as stable as any major BBC activity. It is relevant that the BBC's news website (news.bbc.co.uk) is famous in the Internet world for the stability and longevity of its content: tracking 'related stories' can lead to display of content from as far back as 1997.

Further information:

http://screenresearch.ning.com/profiles/blogs/first-hints-of-bbc-archive

http://www.broadcastnow.co.uk/news/multi-platform/bbc-eyes-online-docs-archive/5005792.article?referrer=RSS

The BBC Charter[18] and associated Agreement[19] detail the constitution of the BBC and its commitments, including the maintenance of an archive, ensuring the archives are kept safely, and giving the public reasonable opportunities to view or listen to archive material.

1.1.2.   *Case study: ORF commercial content sales*

One of the goals of the archive-management at ORF is to earn revenue by selling preserved and migrated content outside our company. Main targets are the common DVD-market, AV-production companies and other broadcasters. Prior attempts to develop and/or join mutual selling-points for "mixed" media-content failed, but the approach is still observed in particular areas. Unfortunately the revenue is far too small to provide a reliable income to finance current and future preservation activities.

Successes

There are some areas were selling TV-content is pleasing and where the revenues are not only covering the costs but producing profit. Unfortunately those areas are mostly connected to common and "old-style" distribution channels; updating those channels is very often expensive, risky and results often in entering highly competitive markets.

DVD/BD-Editions

> ORF is producing editions of TV-programmes on DVD and BD in close cooperation with several Austrian publishing houses. Contents range from fiction-productions (serials, etc.) and "nature-programmes" to Sport-summaries and culture-events. Main challenge is the preparation and acquisition of non-ORF rights as well as the acquisition of licence-add-ons for the intended distribution channel. So exact and accurate planning in the rights-acquisition-phase of each production is vital to be able to benefit from other and/or new distribution-channels after or in addition to the main purpose of the production (which is to be broadcasted).

> Unfortunately, although the above mentioned needs are known and (most of the time) taken account of, the present financial situation during production is often a cause of

---

[18] http://www.bbc.co.uk/bbctrust/assets/files/pdf/about/how_we_govern/charter.pdf
[19] http://www.bbc.co.uk/bbctrust/assets/files/pdf/about/how_we_govern/agreement.pdf

hindrance to do so. In these cases for enabling the rights to publish some productions by new distribution-channels the rights-add-ons has to be bought (of course most times much more expensive than if bought during the original acquisition) or a part of the content needed has to be partly re-produced (for example, the originally broadcasted version of a typical infotainment program contains original shots from another archive or company; but the archive/company does no longer exist and a legal successor can't be identified. So to avoid legal problems, the parts used has to be removed or replaced by similar parts from another collection…; a laborious, time-consuming and expensive process, that can heavily reduce the estimated revenues).

Archive- & Production Sales

ORF is also selling complete programmes and parts of productions as well as raw-footage to other broadcasters and production-companies; since candidates for this distribution-channel are in the majority of cases already foreseen to be distributed, the previously mentioned rights-problem is significantly less frequently. On the other hand the sales at this small market-sector are not producing enough revenue to be a major income-factor for the archive or the whole company.

On demand public sales

To provide our customers with a special service we offer on-demand-copies of wanted items; these are limited to ORF-origin material only to avoid legal problems. Although the acceptance of this service is very high, the time and effort to produce these copies (DVD) is very high and hardly covering the costs.

Although there are concrete plans to offer this service via an online-portal with direct-access to the (parts of the) archive-collection with download and streaming services available, the current legal situation in Austria and the complexity of some rights-situation are the most important causes of hindrance to establish this new workflow.

Unsuccessful approaches

There has always been the hope, that sometimes the income from selling archival content would help to provide enough revenue to provide enough budget for necessary preservation actions and projects. Therefore ORF took part in several projects and trials to test and evaluate different approaches and mechanisms in offering and selling archive-content. Unfortunately non of them were very successful so far, but we're still trying…
Here are now a few examples of those attempts:

AV-media distribution platform (NODAL)

We participated in a project where several different AV-collections tried to set up a concerted distribution-platform for their content. The technical challenges in developing a proper and reliable online-platform were highly successful, the nearly insuperable problem showed up when we tried to develop a price-list for direct-selling via the platform:

In standard licence-trading there are many various aspects and parameters affecting the price; there are those, that can be used to build a "price2use-matrix": intended use, length of licence-period, licence-area, amount of uses (e.g. broadcasts), etc.; on

the other hand there are others, that are way to fluctuant to use in a fixed framework: value of the content (may change in seconds: shots of a certain building may be suddenly be asked for heavily, after it collapsed), relation of the seller to the purchaser (a certain shot may be much more expensive for a competitor (e.g. national private broadcaster) than for a non-competitor (e.g. broadcaster abroad)), volume discounts, etc.

In the end we had to refrain from using the result as a sales outlet, but instead we used it as a platform to promote our content.

MultiMedia-content distribution solution (eCHASE)

We also participated in a project where several different media-collections (broadcaster, publisher, photo-collections, museums) tried to set up a concerted distribution-platform for their content. Again the developing of the technical platform was the easiest part; the challenges and problem were in the parts, where we had to find concerted ways not only to present different content-implementations and formats, but to find common denominators in content. Failing here we ended up in a coexistence instead of a concertation.

Despite these relatively unsuccessful projects, it is still an ORF hope that our participation in projects with a more coherent composition of partners and the so far very positive and pleasant results may lead to some revenue in the future. So we focused more on our core-business ("Birth-of-TV", Videoactive) and joined existing dissemination and access-platforms (e.g. Europeana via EUscreen), estimating to gain interest in our content and therefore support our currently existing and already working distribution-channels

1.1.3.   Case study: INA Inamédiapro and ina.fr

Many of the details of the INA case study below have been extracted from the ITHAKA report "Sustaining Digital Resources: An On-the-Ground View of Projects Today. Ithaka Case Studies in Sustainability"[20]. We have extracted mostly factual information and we encourage the reader to refer to the full report for a wider discussion and also a series of other valuable and relevant case studies.

Overview

The National Audiovisual Institute (Institut National de l'audiovisuel, or INA) possesses a rich collection of French television and radio recordings from the past 70 years. It is among the largest audiovisual archives in the world, with 1.5 million hours of video and sound archives from France's public radio and television stations dating back to the 1940s. Included in the archives alongside the television, radio and newsreel content are substantial collections of photographs, music videos and print materials, ranging from trade magazines and monographs to materials documenting the daily business of television stations, including screenplays and programming grids.

INA provide both commercial content services and public access. The INA marketing and sales division licenses video and audio content to the professional market via

---

[20] http://www.ithaka.org/ithaka-s-r/strategy/ithaka-case-studies-in-sustainability/report/SCA_Ithaka_SustainingDigitalResources_Report.pdf

www.inamediapro.com. The public website www.ina.fr provides to the general public 25,000 hours of content online. While both units create value from INA's archives, they accomplish this in distinct but complementary ways. The commercial licensing group operates as a professional service, using the base of 500,000 digitised hours of content to license content to the professional market, with the goal of increasing sales year to year. The mission of the public website is primarily to provide broad access to a smaller pool of rights-cleared content, and in doing so to build audience and recover costs through a variety of revenue strategies including fee-per-download, customised services and advertising.

Sustainability

The five-year agreement signed by INA and the French government establishing its budget for the period 2005 to 2009 requires that approximately 34% of INA's total budget be generated through earned-revenue activities. In 2008, for example, INA's operating budget of €120 million came from two sources: 66% (€79 million) from the French government via the audiovisual tax (an annual flat fee levied on all households owning a television set), and 34% (€41 million) from a variety of commercial activities.

This means that commercial services are necessarily core part of INA's strategy. Current services include:

- Commercial rights licensing through Inamédiapro: €14.5 million

- Archiving services provided to the public television and radio stations: €12.3 million

- Continuing education courses and degree-granting programmes: €8.2 million

- A production and publication unit focused on the public website www.ina.fr with a variety of revenue streams from the general public, including publication and sale of DVDs produced or co-produced by INA staff, download fees for rental or purchase of video clips, advertising and more: €4.1 million

- A research division licensing INA software and providing consulting services: €1.6 million

While Inamédiapro generates the most revenue (35%) through rights licensing to professionals, www.ina.fr generates the most exposure and broadest access, serving as the public face of the entire organisation. Its contribution to the bottom line is more modest (10%), but its role in developing an audience for INA and as a sandbox for innovation and experimentation to test new revenue models has been very important

Inamediapro

Officially launched in 2004, www.inamediapro.com is INA's client facing interface for archival content, providing user tools and the ability to navigate among the vast holdings of INA's audiovisual archives, including those materials yet to be digitised, as organised and described by INA staff of over 100 archivists. More than a professional sales website, it is the layer that provides a 'window' into the complete body of content for INA staff as well as for external users: all of INA's digitised content is visible through www.inamediapro.com, and the tools that www.inamediapro.com users have access to are the same ones used by

INA staff to navigate the archives. The content is browsable by over 400 thematic topics and searchable by keywords in the metadata, descriptions and other documentation, such as any transcripts that accompany the audiovisual files. In addition, longer video clips can be quickly viewed as storyboards in pre-set time intervals, and can be edited to create shorter clips of just the material that is desired.

The marketing and sales division, responsible for licensing video footage to commercial clients, has a total staff of 67 responsible for client relations (20), back-office sales administration (20), rights licensing (20), marketing (5) and project management (2). In 2008 INA's rights-licensing business processed over 8,000 orders, generating €14.5 million, up from €7 million in 2001. The website www.inamediapro.com has over 8,000 registered users representing over 6,000 companies from France and around the world, about one-third from outside the country. Of these, 800 companies are considered active customers, having placed an order within the past 12 months

About 75% of Inamédiapro customers are companies that use video footage, including broadcasting companies, producers, advertising agencies and internet sites. Another 10% of its customers, mainly television stations, are interested in licensing full programmes, and about 15% are video and music publishers, looking for interesting footage to include in DVDs, 'bonus' DVD clips and DVD/CD packages. At least 10% of annual revenue comes from the international market, and so www.inamediapro.com includes an English-translation feature. The field is competitive, and Inamédiapro management is aware that it must communicate the value of its services to its audience. The launch of the Inamédiapro website in 2004 was a significant event, with advertisements, mailings and banners at the annual industry trade fair for audiovisual content at Cannes (MIP TV). A great deal of care has been taken in developing communications campaigns to describe the features and benefits of the service to its target audience.

Ina.fr

In April 2006, INA launched the 'Archives for All' campaign, unveiling an Open Access site where anyone could view a selection of videos from the archive. Since its launch, the new www.ina.fr website has put 25,000 hours of content online, and 5,000 new hours are to be added each year.

The content consists of approximately 100,000 separate audio and video clips, ranging from a few seconds to over an hour in length – all are rights-cleared videos organised in thematic categories to facilitate browsing. As of November 2008, www.ina.fr was recording about four million page views, 1.5 million videos streamed, and one million visits per month from 600,000 unique visitors. The site serves as the public face of INA and has two principal aims: to build audience by finding ways for people to engage with INA's content, and to generate revenue by monetising both the content itself and the value of INA's audience.

Several types of revenue generation are facilitated by www.ina.fr, including annual advertising revenue of €300,000, revenue from downloading and rental of video clips of €300,000, and as of 2009, a newly created DVD on Demand programme (€25,000 per month to date). In addition, the website includes an e-commerce boutique which sells DVDs, CD and books developed by INA based on the content of the archives. The total revenues that the website generates are approximately €1 million per year. INA has not yet found an economic model to cover the costs of the website. Not counting seven years

of digitisation, or 35 years of metadata creation, or the legal expense of intellectual property documentation, the revenues don't cover current operational costs.

This shows that public access can be hugely successful in terms of reaching an audience and seeing lots of use, but covering the costs of the infrastructure required is not necessarily straight forward even for these ventures.

Combining public access and commercial sales

From INA's earliest days, licensing content has been one of the primary means the organisation has used to support its activities. Since rights licensing is such an important part of INA's revenue strategy, when planning to make some content freely available to the public through www.ina.fr, several controls were put in place to ensure that the offer of free content would not cannibalise Inamédiapro's business of licensing rights to professionals:

- Legally, the public content on www.ina.fr cannot be used for commercial purposes, and those who download a file must agree to terms of sale specifying this restriction

- All images coming from www.ina.fr are digitally watermarked.

- Images on www.ina.fr are in a format of insufficient quality for professional use

- The Open Access www.ina.fr site includes several navigational buttons and menus to re-direct professionals to www.inamediapro.com

Rights protection is a vital element of this system, and it was made possible by the activities of another division at INA. The research and innovation group is tasked with developing new technologies that can advance the quality and efficiency of work at INA and be commercialised outside of the organisation as well. One of its projects is software called 'Signature'[21], which digitally watermarks[22] each licensed image, whether from www.ina.fr or via inamediapro.com, identifying INA as its source of origin. As the stations transmit their broadcasts and INA records them, the Signature software reads the 'fingerprint' of the image, flagging those which may have been improperly used. INA reports that since the software has been in use, it has recouped €500,000 in licensing fees.

There is sufficient differentiation between the free content provided on www.ina.fr and the professional rights-licensing business so that both markets operate without detriment to the other. By devising the Archives for All programme in such a way that it would not threaten the sale of commercial video rights, INA has been able to develop revenue streams on www.ina.fr via pay-per-view and other customisable downloads, while still building the professional market for video programmes and footage. Customised tools and services and access to the full archive directly address the needs of the professional market for video content that Inamédiapro serves.

---

[21] http://www.ina-entreprise.com/to-know-ina/signature.html

[22] Digital watermarking (finger printing) is a complex technical challenge given the need to be robust to transformations in the content (e.g. being able to withstand changes in resolution, encoding, quality, duration etc.). Within PrestoPrime, INA is doing further development work in this area, including fingerprint databases and comparison services. Results of this work will be publicly accessible during year 2 and year 3 of the project.

The www.ina.fr website uses its open content as a way to build both audience and revenue. The free content is editorialised and promoted to attract visitors, which in turn drives advertising and download revenues on the site. By developing partnerships that share INA content via the INA Player on other high-traffic websites, the organisation benefits from an additional channel to increase awareness and drive traffic to the INA site. In turn, the robust activity on www.ina.fr makes it a good place to test a wide range of financial models to see what works best. The team's willingness to experiment with different strategies for both audience-building and revenue-generation has led to some strong new programmes such as DVD on Demand, a viable new revenue stream. In other cases, projects that do not yield the desired results can be abandoned.

Major findings

The conclusions of the Ithaka report are that INA's experience in providing both commercial and public access to content, including the interplay between the two, leads to several key principles:

- The value of content often lies in how it is used, not simply in the content itself.
- A coherent mission with measurable targets helps unify an organisation and keep it focused on achieving its goals.
- Free content can be instrumental in driving revenue.
- Setting up a low-risk environment for experimentation helps foster innovation
- A deep understanding of users is critical to developing new products and services.

1.1.4.   Case study: B&G Images of the Future

Images for the Future is the largest digitisation effort in Europe to date. A consortium of 6 partners (including three archives) is migrating a substantial part of the Dutch audiovisual heritage to a digital environment. The project has three objectives: [1] Safeguarding heritage for future generations. Our audiovisual heritage on analogue carriers has a very short life expectancy. In order to ensure long-term access to our collective memory, it is vital to migrate the material to a digital environment. [2] Creating social- economical value: migrating large quantities of audio-visual material is a precondition to unlock the social and economic potential of the collections. More professionals and individuals than ever before will be able to access material, and new services and businesses will be launched. [3] Innovation: by digitising heritage on a massive scale, a completely new infrastructure will have to be built that can strengthen the knowledge economy of the future. Much is to be expected from recent developments in computer science, especially in the areas of data-mining, information retrieval and in the creation of new environments where content can be used. To achieve these objectives, the cultural heritage sector is challenged to re-evaluate its business models. This article presents the theoretical framework in which Images for the Future is operating and substantiates this framework with concrete examples of results.

Unlocking value

The collections of the Netherlands Institute for Sound and Vision in Hilversum, the National Filmmuseum in Amsterdam, the National Archive in The Hague safeguard the memory of the Netherlands as captured in moving images over the past hundred years. As will be demonstrated below, the collections will represent great social and economic value – if and when they become accessible.

Large-scale migration of analogue audio-visual material for preservation and access is exceptionally expensive. For Images for the Future it involves the selection, restoration, digitisation, encoding and storage of 137.000 hours of video, 20.000 hours of film, 124.000 hours of audio and more than three million photographs. And this is only the beginning. To truly make the material useful, investments will have to be made to enrich the existing metadata so that they can meet up to the requirements for use. And services will have to be developed that make the material meaningful and useful for a variety of user groups. Educational institutions, students, teachers, publishers, television and film makers, web designers, graphic designers, artists, software developers, internet providers, museums, theatres, heritage institutions, libraries, etc. will all profit from content becoming digitally available and accessible over networks.

The total investment of this initiative sums up to 173 million Euro. Not an insignificant sum by any means. A strong business model is necessary to support this kind of investment and prove that such an investment will result in long-term socio-economic returns. To substantiate this assumption, the research institute SEO Economisch Onderzoek executed an analysis of costs and benefits of the Images for the Future project[23]. This analysis followed the so-called "Research into Effects on Infrastructure" framework defined by the Dutch Ministry of Economic Affairs and the Ministry of Transport[24] This framework shows how the costs and benefits of infrastructural investments such as roads, railways etc. can be determined. In addition to 'hard' numbers, environmental and other 'soft' effects are an intrinsic part of the analysis. The socio-economic returns of these structural investments are calculated over a substantial period of time, in the case of Images of the Future, this period was set at 30 years.

The outcome of the Cost-Benefit analysis was positive: "The total balance of costs and returns of restoring, preserving and digitising audio-visual material (excluding costs of tax payments) will be between: 20+ and 60+ million." The "plus" represents non-quantifiable indirect effects and social benefits that go beyond the users' benefits. Direct effects concern the market for culture and heritage, and the market for information. The project is financed through the so-called Economic Structure Enhancement Fund. This fund is used to finance large-scale infrastructural investments. Direct effects of the investment are revenues from sales, access for specific user groups, the repartition of copyright for the use of the material and so on. The indirect effects concern the product markets and labour market. Within the duration of the project (2007-2014), Images for the Future will generate an equivalent of 300 man-years of work. Social benefits take place outside the markets. For Images for the Future these benefits involve effects like the conservation of culture, reinforcement of cultural awareness, reinforcement of democracy through the accessibility

---

[23] Images for the Future: Outline of benefits. Amsterdam: SEO Economisch Onderzoek.
[24] Eigenraam, C., C. Koopmans, et al. (2000) Evaluatie van infrastructuurprojecten; leidraad voor kosten-batenanalyse, Deel I: Hoofdrapport & Deel II: Capita Selecta. Den Haag: Centraal Planbureau.

of information, increase in multimedia literacy and contribution to the Lisbon goals set by the EU.

It's important to note that much of the material kept by the archives is copyrighted. Images for the Future collaborates closely with this important stakeholder group to ensure that the material is used in accordance with predefined arrangements. The underlying assumption is: "everybody wins"; material is made accessible to end users in a meaningful way, and rights holders receive their fair share of the income generated.

Although some of the predicted economic effects are debated, this analysis is in line with other prominent studies on the subject of the effects of digitisation, most recently demonstrated in a study on the effects of file sharing in the entertainment industry. (Huygen 2009) Large private enterprises such as Google have signed off on business plans underlying the same assumption. (Varian 2009) Most authors agree that while the new business models have yet to prove durable, it remains crucial to create an environment in which our heritage becomes part of the modern intellectual discourse and that it will create economic value in some form or other.

Business Model

The question remains *how* to maximize the socio- economic impact of the project. When examining the potential positive return of an investment, it is useful to have a tested model as a starting point. In this case we use the business model canvas developed by Osterwalder and Pigneur (Figure 2), that has become a popular reference point. (Osterwalder 2003)
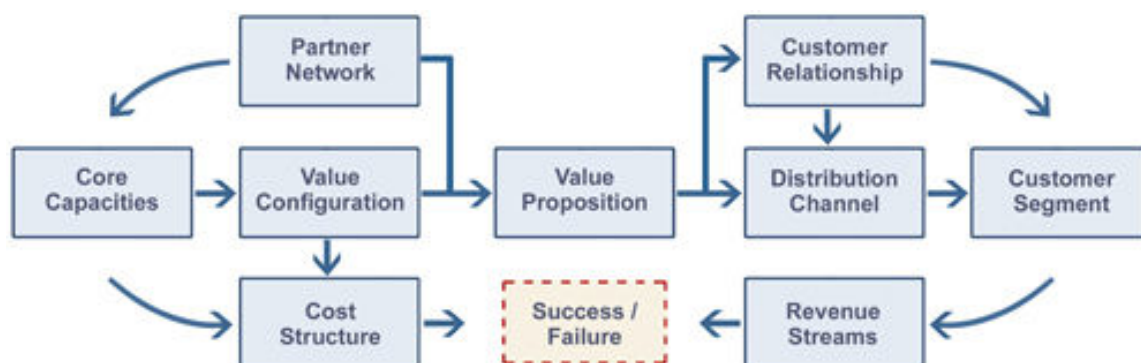

Figure 2 Business Model Canvas. Osterwalder/Pigneur 2003

In this context a business model is based upon the logic of an organization to achieve its (financial) goals, using nine interrelated building blocks:

The value proposition of what is offered to the market;

The segment(s) of clients that are addressed by the value proposition;

The communication and distribution channels to reach clients and offer them the value proposition;

The relationships established with clients;

The key resources needed to make the business model possible;

The key activities necessary to implement the business model;

The key partners and their motivations to participate in the business model;

The revenue streams generated by the business model (constituting the revenue model);

The cost structure resulting from the business model. (Osterwalder 2003)

It is important to note that – using this model – not only the revenue sources have to be taken into account, but also relationships between the different building blocks that will change as one of the elements changes.

In our case the value proposition of the archive changes while artefacts are being migrated from analogue carriers of information to digital assets. With this migration, the variable cost of distribution has decreased tremendously, which in turn opens up ways to reach many new client segments that were previously difficult or very expensive to reach. Whereas the main client segment of organizations in the cultural heritage domain used to be limited to professional researchers and program makers who would come to the physical location to conduct research, a larger public can now be engaged in using (parts of) the archive that have a specific interest to them.

Value Proposition

The value proposition of the various services provided can be positioned in terms of a model of types of service use (accumulation, construction and archiving) as shown in Figure 3.
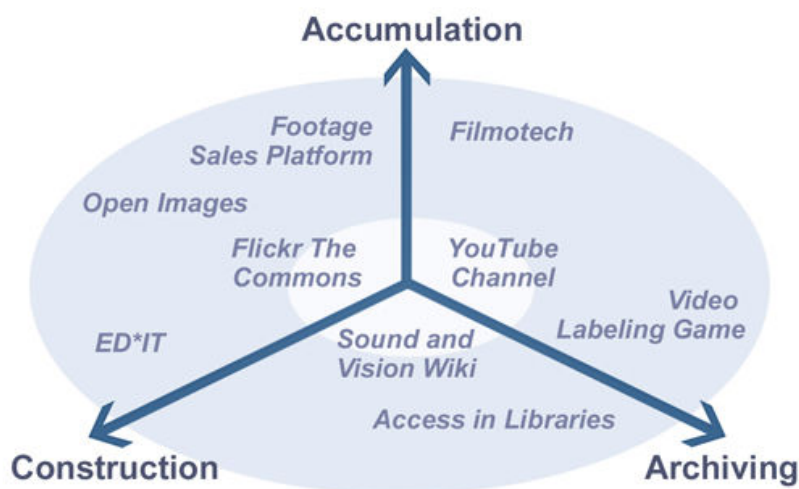

Figure 3 Accumulation, construction and archiving model for services provided by Sound and Vision.

Service	Description	Value proposition	Creation
ED*IT	ED*IT (http://www.ed-it.nu/) is a multi-medial educational platform that can be fully integrated in existing learning environments. The service will be used for the first time in the 2009-2010 school year and will offer content from Sound and Vision, the Filmmuseum, Public Broadcasters, Nationaal Archief (the National Archive of the Netherlands), and four museums within the fields of technology, national history, history and ethnography.	Access to an elementary collection is free, schools pay a fee for added services on the platform (edit facilities, lesson templates)	Construction Accumulation
Open Images	Sound and Vision and Knowledgeland are developing Open Images (http://www.openimages.org). The aim of this platform is to offer online access to a selection of	Allowing access to material in near-broadcast quality for	Accumulation Construction

	archive material to stimulate creative reuse. Reuse includes remixing of archive footage in new videos.	reuse in new productions. Facilitate orders of high-quality material. Offer snippets from DVDs as teaser.	
Sound and Vision Wiki	The aim of the Sound and Vision Wiki (http://www.beeldengeluidwiki.nl) is to add contextual information to the collection of Sound and Vision by inviting "The Commons" to document and share their knowledge about programs, media personalities and media related topics. This is not restricted to universities or journalists; a broad constituency of enthusiasts should be encouraged to contribute to their heart's content.	Using outside knowledge to secure knowledge on the collection and by doing so, increasing accessibility and reuse.	Construction Archiving
Nationaal Archief on Flickr The Commons	Within the project Images for the Future, a lot of the items in the collections (in this case from the Nationaal Archief), lack adequate metadata Given the vast quantity of stills that need to be annotated, the Nationaal Archief realised they needed a broader public to help them and decided to join The Commons initiative. The Commons was launched early 2008 as a collaboration between the Library of Congress (LoC) and Flickr.	Using outside knowledge to secure knowledge on the collection. Using an existing platform (plus active community) as a proportional tool.	Construction Accumulation Archiving
Video Labelling Game	To turn the incentive "play and competition" into the foundation of a new service, Sound and Vision has taken the initiative to develop a video tagging game, in close collaboration with software developer Q42, KRO broadcasting and VU University Amsterdam	Using outside knowledge to secure knowledge on the collection in a game-like setting. Use space for advertising.	Archiving Accumulation
YouTube Branded Channel	In 2007, Sound and Vision was one of the first archives in Europe to offer access to a selection of their collection though a branded channel on YouTube. Currently, over 200 items can be viewed online, most of them originate from the newsreel collection. The clips have been viewed a stunning 1,2 million times, making the YouTube branded channel one of the most visited in The Netherlands.	Using an existing platform (plus active community) as a proportional tool.	Archiving Accumulation Construction
Footage Sales Platform	The "creative industry" is an umbrella term for companies that include publishers, television and film producers, distributors, broadcasters, internet providers and so on. Next to the education sector and the general audience, the creative industry is the third target audience for Images for the Future. Services are still in the developmental stage, but will include an online footage sales platform, aimed at the international professional market,	Offer access to a part of the collection to professionals in the creative industry. Fee based on type of reuse and intended audience.	Accumulation (by professionals)
Filmotech.nl	"Filmotech.nl" is the working title of a soon to be released video-on-demand platform. Initiated by the Association of Dutch Feature Film Producers (NVS), Filmotech.nl is a collaborative effort of the Dutch film industry, including three partners from Images for the Future: Sound and Vision, the Filmmuseum and Knowledgeland. The platform aims at providing access to the complete Dutch film repertoire and to a substantial selection of European content: feature films, television dramas, shorts, documentaries and animation	Offer access to a part of the collection in a platform aimed (at least initially) at the consumer market. Income is generated on a pay-per-view basis.	Accumulation Archiving
Access though libraries	negotiations are ongoing between the Images for the Future consortium and rights holders about the possibility to offer access to the collections in libraries through existing services operated by the Netherlands Public Library Association (VOB), notably its common catalogue Bibliotheek.nl.	Negotiate a fee with copyright holders concerning the availability of content in libraries. Charge for access through lending fees or subscriptions. Using Web2.0 to enhance search experience.	Archiving Construction

Summary

A myriad of services can be created as cultural heritage splashes into the digital age. Due to user-demand, growing availability of digital collections, increased connectivity, maturing of data mining, and semantic web technology (to name just a few), thinking about services is evolving in an increasingly rapid pace. There is clear interest for well-documented and tested services that add value to digitised collections that can serve as best practices. As new value propositions will be key in giving organizations the perspective, the confidence, the contacts, the tools and the funding to launch major digitisation projects[25]. In this case study, the notion of service development by cultural heritage institutions has been linked to value creation by users. Perhaps not surprisingly, business models can be based on a variety of value creations, from subscription-based models (i.e. ED*IT) to completely free services (i.e. Open Images). This exercise proved to be insightful. It provides proof that digitised cultural heritage can be used in endless ways and within various contexts, meeting expectations of a heterogeneous group of users. Initial findings are encouraging. In the Netherlands for example, organizations working within Images for the Future have recently witnessed a steep rise in requests by professionals, consumer sales and the utilization of material in education. Also, public-private partnerships are being forged that will certainly increase the impact of the project through (for example) the exploitation over new distribution channels and devices.

1.2.   *Value-chains and business models for preserving file-based content*

In this section we look at new and emerging business models and value chains for preserving files. The main and recent developments here seem to be using the use of service providers for hosting the archive. This approach is very much in its infancy, with the vast majority of archive hosting currently done in house.

Our focus is on file-based assets: digitisation, transfer of digital video, and preservation of analogue material were all covered by the preceding PrestoSpace project.

Overall, the idea of 'archives as services' from third parties is very much a new concept in the audiovisual archiving community.

There are existing commercial services for temporarily hosting AV content so it can be distributed, e.g. VIIA from Ascent Media[26] but fewer if any that provide large-scale long-term preservation with online access to the content.

There are existing services, e.g. Iron Mountain[27], that have existed for some time to provide a 'tapes in a bunker' model for a deep archiving, i.e. a safety copy, but these services have until now provided very little in the way of online access. They are just beginning to move in this direction.

---

[25] Blue Ribbon Task Force Interim Report.:Sustaining the Digital Investment. consulted January 16, 2009. http://us.ft.com/ftgateway/superpage.ft?news_id=fto072720071832456749&page=2
[26] http://www.ascentmedia.com/
[27] http://www.ironmountain.com/index.asp

There are also some emerging activities in the AV preservation community, e.g. AVAN[28] (Audiovisual Archive Network) in the States that is exploring hosted digital preservation services and the ProArchive[29] service from the Sound and Vision in the Netherlands, but these are currently few and far between. Both are described in more detail below as cases studies.

There are also some embryonic services in the digital library space, e.g. DuraCloud[30] that are looking at using 'utility storage' models for long-term preservation of files, but these aren't specific to AV. It is not clear that they can handle with the large file sizes and data volumes for AV archives. OCLC is another example[31] of this type of service.

At the same time, the cloud storage space, including archiving is becoming a very hot topic, with some new big players entering the area alongside the more well known and established cloud players, e.g. Amazon. These include EMC[32] and IBM[33].

On the one hand, storage as a service, especially cloud storage, is now maturing to the point where people are starting to use it (outside of the AV sector, e.g. for business data). On the other hand, archiving as a service, especially for AV, is very much a new concept and is very underdeveloped. The two things go together quite naturally given the economies of scale achieved by large storage services which can help lower the cost of long-term preservation whilst also supporting online access. Therefore, convergence can be expected and new breeds of cloud-based archive as a service players may well emerge.

1.2.1.   Case Study: ProArchive

In many different places in the Netherlands, films, video productions and sound recordings are stored featuring the development of towns, regions, society and politics, or even entire business branches and the (in some cases obsolete) production processes and professions associated with them. Without all this visual information, the history of the twentieth century would be incomplete and hard to understand. This is why AV content is often used for research and educational purposes, in film, radio and TV productions, at exhibitions or simply for a walk down memory lane. In short, these audiovisual sources form a natural part of our national cultural heritage.

Preservation backlogs of AV carriers occur in the entire heritage sector. The reason is that AV carriers are more vulnerable than our traditional media, and extremely expensive to preserve. We cannot process this backlog with the current analogue preservation method: the costs are too high, and moreover, in ten years or so the preserved collections will need to be preserved again in the same way – at the same high costs. This makes the shift towards digitising audiovisual sources necessary.

---

[28] http://www.archivenetwork.org/
[29] http://instituut.beeldengeluid.nl/index.aspx?ChapterID=8633
[30] http://duraspace.org/duracloud.php
[31] http://www.oclc.org/us/en/digitalarchive/default.htm
[32] http://www.theregister.co.uk/2009/11/03/cisco_emc_acadia/
[33] http://www.theregister.co.uk/2009/10/06/ibm_xiv_based_cloud/

Since 2006 Sound and Vision has received support from the Ministry of Education, Culture and Science to extend its digital infrastructure to support other (and primarily smaller) AV custodians in the broadcast (regional broadcasters) and heritage domains.

This resulted in the PROARCHIVE service that aims at regional broadcasters and other organisations that wish to manage and disclose their audiovisual archives in a sustainable and professional manner. As Sound and Vision has already developed the digitisation structure, storage, and disclosure of AV material, the costs will be relatively low.

PROARCHIVE finished as a pilot in 2009. Sound and Vision will continue to build the service in 2010 specifically for regional broadcasters. The aim is to have the full service and cost-models specified and the service to be able to handle regular file ingest from two or three regional broadcasters by the end of 2010. Structural costs will have to be covered by future additional funding and will be considerable when the service is eventually to handle materials from 13 regional broadcasters. PROARCHIVE, therefore, is still running at experimental level at least for 2010.

Using the PROARCHIVE infrastructure would mean that regional broadcasters will be able to search, view, order, exchange and manage AV content (entire programmes as well as clips) in an easy way. The metadata editor allows to add metadata and the rights manager enables users of the service to decide whether or not their content may be reused, and if so, under which terms and conditions. In other words, they will remain in control of your own collection.

PROARCHIVE's digital infrastructure will make it very easy to publish and exchange collection(s) - or parts of it – with other parties through various Dutch AV platforms, such as Academia and Teleblik. With broadcasting companies, for instance, teaching establishments, AV producers, or Sound and Vision and its incredibly extensive multimedia catalogue that offers a junction for exchanging and broadcasting (archive) material.

Besides the PROARCHIVE service to be build for use by regional broadcasters, Sound and Vision is also investigating possibilities for offering (parts) of its digital infrastructure for use by heritage institutions, including museums, archives and libraries, that conserve AV collections. Sound and Vision has set out to inventory as many user experiences as possible in 2008 and 2009 by running pilots for a wide range of establishments with a wide range of uses. Pilots ran not only for regional broadcasters but also for non-broadcast organisations, such as a library in Zeeland, the Volkskrant national newspaper's TV, the Film Museum in Amsterdam, and the Erasmus Medical Centre in Rotterdam

Selling points

Eventually, Sound and Vision aims to extend its digital infrastructure with services and tools for other organisations, with the following benefits in mind:

- The system can accommodate both large and small collections;
- All content is easy to find via an extensive metadata model;
- AV content is stored using standard file-formats, e.g. MXF D10/50.
- One can utilize Sound and Vision's knowledge and expertise in the field of AV archiving;

- One may access and exchange content with cultural heritage and other educational establishments, organisations that manage archives, broadcasters, and the creative industry;
- Sound and Vision can offer digital sustainability due to continuous developments and investments in new, future-oriented system architecture;
- One will hardly need to invest in hardware, software and infrastructure.

1.2.2.   Case Study: Audio Visual Archive Network (AVAN)

Note: much of the information below has been extracted directly from the AVAN website and is reproduced here verbatim. For the full text and further details see the AVAN website (http://www.archivenetwork.org/)

Overview

The Audiovisual Archive Network (AVAN) is an independent, non-profit organization dedicated to preserving and providing access to audiovisual content. AVAN is founded and operated by archivists with deep understandings of preservation, access, technology, and metadata. There is no existing service specifically dedicated to providing clip-level access to primary source sound and moving image material, nor to help archives and creators preserve their digital masters. AVAN has been created as one solution to this problem.

Services

"AVAN offers two core complementary services to help archives achieve their preservation and access mandates."

- "The aggregated Library makes it possible for scholars, students, and the public to discover archival time-based content using cutting-edge tools that no single archive could afford to create on their own."

- "The Digital Repository Service preserves digital files for archives and creators lacking infrastructure to manage digital preservation efforts themselves, ensuring that the content will be available for future generations."

This model is not new in the cultural heritage space where 'browse' quality media is made accessible online for public access and education, supported by services for the long-term preservation of master copies.

Target Market

"Preservation of analogue AV content is now accepted to require transfer into the digital domain, which creates 'reborn' digital files. "

"Since many archives do not have the infrastructure to preserve "reborn" digital files, much important content on magnetic media is not being encoded and is therefore being lost. AVAN will offer a solution to these archives' lack of infrastructure, which will aid the archives' preservation efforts. Improving archives' digital preservation efforts will result in more content made available to researchers, scholars, and the public."

AVAN is clearly targeting the small to medium sized archives who are moving into the digital domain and do not have the expertise or infrastructure to effectively preserve their contents as digital files.

Charging Model

"AVAN has two main services: an online Library to access the content, and a Digital Repository Service (DRS) where archives can store and preserve their digital files. The two services are not interdependent: archives can make use of the Library, the DRS, or both."

"There is no charge to an archive to contribute content to the AVAN Library. Archives retain full ownership and control the access rights to their content. They choose where their content is made available, even at the item level. Copyright and donor requirements might influence an archive's decision in allowing an item to be available on the open internet (the public site), downloaded, or only streamed and used within AVAN's educational site. Archives might choose to place the same content in both the public and subscription sites. They can even choose in which subscription area an item can be accessed. For example, an archive might feel that some content is not appropriate for K-12, so they could exclude the item from K-12 access."

"AVAN's second core service is the Digital Repository Service (DRS), a digital repository where non-profit organizations and individual creators can store and have their digital masters preserved by a low-cost service dedicated to time-based content. Files are verified and authenticated. Archives are given individual portals where their authorized staff can add technical, preservation, and rights data and view surrogates of the digital objects. All data is retained in the DRS database, as well as stored with descriptive data and the digital object "essence" file (e.g., the actual audio or video content file) in an open MXF wrapper. (AVAN will create an open archival MXF specification as part of the prototype project, and share it with the community.) Metadata and files are not locked into proprietary software, so they can be exported for archives upon request."

So, in summary, lower resolution content for public access and educational use can be stored at no charge using AVAN. However, preservation and storage of master quality material is a commercial service at a cost that is currently not determined.

Business Model

The business model for AVAN itself seems relatively clear. Subscribers pay for access to content for use in education and these subscription fees cover the costs of online hosting of archive content at relatively low resolutions. This removes the need for archives to pay to make their contents available online, which in turn means the barriers to entry are as low as possible in order to stimulate content contribution and the creation of a critical mass of content within the AVAN service.

For the master quality content, the archives are charged for the preservation service. It is not clear whether the costs of operating the service by AVAN are fully passed on to the customers, or whether they are subsidised by either core funding or the revenues from the subscribers to the educational library.

AVAN started life with support from the Andrew Mellon Foundation to develop a business plan and since then has secured funds ($720,000) for a pilot phase, and is seeking further investment. In 2009, AVAN was incorporated as a non-profit organization in California, and is designated a 501(c)3 tax-exempt organization by the IRS.

Given the support from the Mellon Foundation and the services provided for educational use of content, it will be interesting to see how the business model pans out in comparison to other similar ventures such as ArtStor that operate in this space.

Current Status

AVAN is planning a prototype of its services involving a number of contributors who will provide a representative, but relative small number, of items. Partners in the prototype are both production organisations[34] for example local radio and TV stations, as well as University organisations[35] for example libraries and media archives. The prototype is planned to start in Spring 2010. When the services will be fully operational and open outside the prototype partners is not yet clear.

1.3.   Key Findings and Conclusions

Preservation, Access and Sustainability are all interconnected and a strategy for one can't be developed without considering the others.

The strategies for preserving content (e.g. which formats to use, or whether to do the work in house of use a service provider) need to be considered in the context of who needs to access and use the content and the economics and business models of long-term sustainability.

Sustainable preservation through online commercial sales or public access is not an easy business model to make work in practice

There are many people who have tried and failed to use content licensing as a way to generate revenue to sustain preservation. More successful are those who combine commercial licensing with public access services. This typically aligns with an archive's core mission of providing access to its content whilst at the same time recognising that monetisation can bring in complementary and additional revenue streams.

Digital preservation and archive hosting services are in their infancy, but are starting to emerge. This area can be expected to grow quickly over the next few years.

Third-party services for long-term hosting and access to digital AV assets are currently very few and far between. However, for many small to medium archives who are ill equipped to cope in a world of digital files, the outsourcing of preservation to third-parties is an attractive, although currently unproven model. Beyond the issues of trust (see D2.3.1), the economics of remote data centres (power, cooling, economies of scale, network accessibility) along with the drivers for immediate online access to archive content are likely to come together to provide strong drivers for more widespread uptake of digital archives as services.

---

[34] http://www.archivenetwork.org/prototypecontentpartners%3Aproductionorgan
[35] http://www.archivenetwork.org/prototypecontentpartners%3Aproductionorgan2

2. Comparison of preservation strategies

This section examines various specific strategies for audiovisual preservation. These include long-lived media, multivalent as a long-lived decoder, use of file-format migration, and maintenance of software/hardware codecs.

There are two main areas in this section:

1. Preservation at the AV content level. This concerns file formats (e.g. audio/video encoding and wrapping), the use of format migration (when, how, what to migrate to) and also multivalent as a complementary approach where a virtual machine is migrated instead of the file format.

2. Preservation at the bit level. This concerns the use of mass storage technology (tape robots, disk servers) for audio/video files, or alternatively adopting 'long lived media' using an 'items on shelves' approach. Questions addressed here are what are the long-term costs of using these types of storage strategy and in which circumstances is one preferable to another.

These two areas aren't independent, for example the decision on when to migrate to an uncompressed format is highly dependent on the affordability of the extra storage space that this needs. These specific areas are also linked to overall preservation strategy. For example, if content is frequently accessed and needs to be regularly transcoded to contemporary delivery formats then this has a major impact on the formats and storage strategy.

This section also includes an examination of cost. This forms the basis of comparing and assessing preservation strategies. Cost isn't just about technology, but is the total cost of ownership (people, facilities, power, cooling, space, licensing, support, maintenance etc.) and how this can be estimated over the long-term.

The process of comparing preservation strategies and building a plan for preserving objects using one or more of these strategies, e.g. format migration or emulation, is referred to as 'preservation planning' (see the definition by the PLANETS[36] project).

The PLANETS project has done extensive work on a general framework for preservation planning, including development of the PLATO[37] tool to support the planning process. A recent and in-depth review of the activities involved in planning, including how to evaluate candidate preservation strategies is provided by Becker et al[38] with further useful resources on preservation planning and strategy available from the PLANETS website. In this way, we would suggest that one approach to preservation planning would be to use the specific strategies considered below within the PLANETS framework.

---

[36] http://www.planets-project.eu/
[37] http://www.ifs.tuwien.ac.at/dp/plato/intro.html
[38] http://www.ifs.tuwien.ac.at/~becker/pubs/becker-ijdl2009.pdf

2.1. *Preservation through Long-lived Media*

Audiovisual media are said to be short-lived, but nothing is a short-lived as technology in the IT industry: disks that need replacing every three to five years[39], data-tape formats that are superseded roughly every two years with players that are only backwards compatible for two generations as is the case for the LTO data tape roadmap[40].

In comparison, there is film from 100 years ago that can still be played today. Of course it would be seen as too fragile, so this comparison isn't strictly true. The film would need to be copied, very carefully, and that copy could then be played. The same is true for wax cylinders, except they are even more fragile – and both film and wax are susceptible to mould and various other forms of biological and chemical attack. Shellac audio recordings, and vinyl, can also last a century or more. There are many examples of 78 rpm recordings from the first decade of the 20th century in the British Library Sound Archive[41]. Again, these are too fragile and precious to be used as actual 'user copies' – but they do prove century-long survival, and modern playback technology works (with relatively minor adaptation or adjustment) on century-old film and audio formats.

The consequence of the disparity between surviving technology from the 19th century vs. the 'continuous obsolescence' we now experience in the IT industry, motivates the wish to find some more permanent approach to the use of modern technology.

A basic issue is cost: cheaper technology doesn't last as long, and this has been true since clay tablets were superseded by papyrus. In fact a general trend can be deduced:

Medium	Storage Density bits/cm²	Life, years
Stone	10	10 000
Paper	10^4	1000
Film	10^7	100
Disc	10^{10}	10

Table 1 Life expectancy of storage media[42]

Table 1 supports the generalisation that each change to a new technology produces a result that is 1000 times more compact (which implies a cost reduction), but lasts 1/10th as long. We could speculate about another column in the table: the time between such changes – which would perhaps be roughly the same numbers as for 'Life, years'. If that trend continues, we're due – now – for a technology that has a 10^{13} bits/cm² density, but would only last a year! Would we want such technology?

In fact we have 10^{13} bits (and greater) technology for semiconductors – but at a much higher 'per bit' cost than for magnetic storage (disc or tape). The advancement of data tape also continues, with 35TB on a single reel of tape now possible[43]. Holographic

---

[39] Schroeder, B. & Gibson, G. A. (2007). Disk failures in the real world: What does an MTTF of 1,000,000 hours mean to you? *5th USENIX Conference on File and Storage Technologies*. February 2007, from www.usenix.org/event/fast07/tech/schroeder/schroeder_html/
[40] http://www.ultrium.com/technology/uroad.php?section=0&subsec=uroad
[41] http://sounds.bl.uk/
[42] R Wright, In a post-media world – what *are* we keeping and where are we keeping it? MIT6, April 2009
[43] http://www.crunchgear.com/2010/01/22/fujifilm-and-ibm-develop-magnetic-tape-that-stores-35tb/

storage[44] also has potential for 'tens of terabytes' (meaning 10^{14} bits), but per cm^3, not cm². Holographic storage is also potentially more stable than is magnetic-based storage. The stability of the companies developing it is another matter, e.g. as illustrated by the recent demise of InPhase[45]. This uncertainty of the lifetime of the vendor as opposed to the lifetime of the technology is not an unfamiliar problem, e.g. with Plasmon and their 50 year Magneto-Optical devices as another example[46].

Is a longer-lived storage method a practical possibility, for audiovisual content? Clearly film still exists (though the continued production of film stock through the 21st century is by no means guaranteed[47]).

There are three major possibilities for century-length storage:

1. analogue film (and by analogy vinyl recordings for audio)

2. 'digital film'

3. a change to the method of access to digital storage, to somehow standardise (and maintain) an access method that would remain viable despite the rapid changes in the IT industry

In this section we concentrate on digital solutions.

2.1.1.   Digital Film

There is now a technology to write, using lasers, high-density digital information onto microfilm or other film formats.

Digital images for eternity : colour microfilm as archival medium
 http://cat.inist.fr/?aModele=afficheN&cpsidt=18918957

Long-term preservation of digital images on colour microfilm (2005)
http://en.scientificcommons.org/20358913

The main problem with assessing the practicality of 'emerging technology' is a realistic cost estimate. If personal computers were not manufactured in millions of units, their price could well be 100 times higher. Using a laser to write "bits" to high resolution, ultra-stable, film – and preserve an analogue colour image at the same time and on the same piece of film – is attractive for image preservation at the high-value, low-volume end. It might convince the curator of images at a national picture gallery, for instance – who might have something like 10,000 to 100,000 images of very high value, representing 'master photographs' of the entire collection. That collection could well be the result a century or more of effort and investment, covering the work of dozens of people and funding in the hundreds of millions (of Euros, pounds or anything similar).

---

[44] http://en.wikipedia.org/wiki/Holographic_data_storage
[45] http://televisionbroadcast.com/article/94340
[46] http://www.theregister.co.uk/2008/10/08/plasmon_in_administration/
[47] See http://motion.kodak.com/US/en/motion/Products/Chronology_Of_Film/chrono4.htm for the trend at Kodak, the main producer worldwide: mainly discontinued film formats since 2005

Television has 25 images (frames) per second (30 in North America), meaning 90,000 per hour. Clearly there are very different situations when comparing 100,000 images representing the best possible photographic record of a major cultural institution with 90,000 images for one hour of television.

Commercial costs are not available, but Fraunhofer have worked on use of this technology for audio[48] They demonstrated a storage density of 4 KB/mm^2, which would be 400 KB/cm^2, or about 4 Mbits/cm^2. That's still in the range of the 10^7 bits/cm^2 that has been the effective optical resolution of film for over a century.

The advantages of 'bits on film' are two:

1. the reading-back is an optical process, and so could be re-constructed in a century with no particular difficulty. Either technology similar to, or much better than, current optical technology will exist then – or there will be some sort of post-apocalyptic loss of technology in which case people will have other things to worry about.
2. the material is stable. If left is a closed box, in reasonable humidity and temperature conditions, the life expectancy is centuries.

The problem is cost. As stated, until there would be wide take-up the technology would remain expensive. The cost of preparing a single microfilm frame at current rates is of the order of one US dollar. That would be about 1 MB, which is about 5,000 times more expensive than magnetic storage (USB hard drives of 1/2 TB are roughly $100 as of Autumn 2009). However the USB hard drive would last a few years, and the microfilm could last 500 years.

The real problem any archive would have, would be finding anyone in the accountancy department who would be prepared to fund anything based on a 500 year amortisation.

2.1.2.   Changing the Method of Access

We already have examples of long-lived digital media:
- magneto-optical discs (Plasmon) should have had a century of life. Plasmon certainly advertised a projected minimum 50-year – but the company has now folded, so the devices to read those discs are now a severe problem;
- variants of CD and DVD technology using etched bits rather than change in dye colour, and using glass or other non-scratch protective layers, for decade or century (or more) permanence.

But, as with the Plasmon M-O discs, will there be devices to read the data, and software to interpret it properly? The short answer is "no", based on past experience: 8" and 5¼" and 3½" floppy discs; a long list of defunct data tape formats; high-density alternatives to floppies, like ZIP discs and miniature data tape devices (on equally obsolete computers, like the ICL-Sinclair OPD with the microdrive[49]).

---

[48] DANOK: Long Term Migration Free Storage of Digital Audio Data on Microfilm;
http://www.imaging.org/ScriptContent/store/epub.cfm?abstrid=38887
http://publica.fraunhofer.de/dokumente/N-94186.html
[49] http://www.rwapsoftware.co.uk/oneperdesk/technical_info.html

The missing link is an access method that would remain viable despite the rapid changes in the IT industry. That access method would have to be standardised (and maintained). Clearly the proprietary route is a trail of ex-companies, and so the alternative route could be an open-source technology for storage access, that was somehow independent of the whole storage stack of any IT system.

It is possible to imagine candidates. For example, there is an open standard for the digital representation of video – and its associated audio – which is independent of all file systems, because it preceded file systems and doesn't use them: SDI, the SMPTE Rec 601 for digital video, which defines a real-time bit stream for 'carrying' a TV signal as a stream of bits instead of as an analogue signal.

Anyone today who wants to connect a computer to SDI needs an SDI card, and software to drive it. The software comes with the card, and 'pulls in the bits' from the SDI signal (connected via BNC connectors, which have been around for 50 years and show no sign of going away) and hands these to the computer operating system as some sort of file.

Writing an SDI sequence to storage eliminates file-system dependence. Anyone interested in reading the stored data in 50 years needs:

1) an up-to-date equivalent of an SDI card;
2) a device that reads the actual storage.

Point 2) is the difficulty, unless we assume that something like a DVD reader will be available in 50 years, and so can blithely purchase (at considerable extra expense) super-high-quality DVD-type storage media now, and cross our collective fingers.

One possibility is the development of flash technology, which has no moving parts and so needs very little in the way of a read-back device. Solid-state storage today is packaged in a range of formats, but not so wide a range: devices costing under £20 are readily available that read all the current common form factors.

When and if storage with no moving parts becomes cost effective compared to magnetic storage (see cost section later in this report), then we could then easily conceive of putting SDI onto such storage, and expecting in 50 years that hardware to 'access the bits' and transmit them as SDI could be easily constructed, and the equivalent of today's SDI card could also be easily constructed.

There appears to be a lack of data on long term data retention for flash memory. There have been initial concerns over charge leakage, especially for multi level cell memory, but the research[50] done here was over 10 years ago. There have not been widespread reports of data retention problems with flash (not in the same way as there has been for spinning disk for example). Manufacturers claim 10 year data lifetime or higher, with some specific write once technologies (WORM for solid state memory) claiming 100

---

[50] Rajiv Verma. Flash memory quality and reliability issues. Memory Technology, Design and Testing, 1996. Records of the 1996 IEEE International Workshop on, pages 32–36, 1996.

Seiichi Aritome, Riichiro Shirota, Gertjan Hemink, Tetsuo Endoh, and Fujio Masuoka. Reliability issues of flash memory cells. In IEEE, volume 81, pages 776–788, May 1993.

years[51] data lifetime. There is not the evidence to suggest that these lifetimes will not be achieved.

2.2.   Preservation using Migration

2.2.1.   Video File Format Migration

It is now standard preservation practice[52] to store audio data uncompressed, e.g. as Broadcast WAV (BWF)[53] files.

Unfortunately, for video, the data volumes and bit rates for uncompressed content make this prohibitive. This typically leads to the use of compression, both in the production process and then for long-term preservation.

The development of new encoding schemes to compress video more efficiently takes place rapidly, for example witness the use of the relatively new H264/MPEG-4 scheme for HD content. This in turn means that these encoding schemes are at risk of equally rapid obsolescence.

At the same time, the cost of storage continues to fall rapidly, mirroring Moore's law for computing power[54] with the volume of storage per unit of media (e.g. a hard drive) doubling roughly 18 months. Given the unit price for media remains roughly constant, the law can be formulated another way: you get 100 times the storage for your money every decade.

The risk of obsolescence and the falling cost of storage means that a obvious preservation strategy is to keep content in its original production format for as long as sensible and then move to an uncompressed format when storage is affordable.

The question is what process to follow to assess what video encoding file formats to use and when, how to decide whether it makes sense to use a half-way stage of lossless compression, and how to decide if alternatives to file format migration make sense, e.g. maintaining decoders for original file formats or using multivalent techniques.

This section models preservation of video essence where migration takes place between file formats. It is assumed that any corresponding audio will be stored uncompressed from the outset and hence no format migrations are needed. The document then extends the model to include multivalent and other approaches to show how they fit into an overall preservation strategy.

This section doesn't model wrapping, unwrapping or transwrapping (e.g. use of MXF). A separate model is needed for this. It is important to separate wrapping from essence encoding as the two activities are not the same (they involve different formats, resources, timescales, objectives etc.) and can be done independently in the digital domain.

The model presented here doesn't include the creation and use of 'access' copies, e.g. lower resolution versions used in browsing, viewing etc... Preservation and access copies

---

[51] http://www.sandisk.com/business-solutions/archiving/sd-worm
[52] for example see IASA TC-O4 http://www.iasa-web.org/special_publications.asp
[53] http://en.wikipedia.org/wiki/Broadcast_Wave_Format
[54] http://en.wikipedia.org/wiki/Moores_law

have different purposes, formats, lifetimes, storage requirements etc so should be treated separately.

Use cases

The model presented in this section is designed to cover the following preservation use cases.

- IMX to MXF. This is a RAI preservation scenario where digital video tape in IMX format needs to be transferred to MXF wrapped MPEG2 (SD video in D10 standard). This is a transfer from a proprietary video format (Sony IMX) to a open file-format without any change to the encoding of the video.

- D3 to uncompressed files. This is a BBC preservation scenario. D3 tapes are transferred to uncompressed files by taking the AV bitstream from an SDI feed and then storing this in an MXF wrapper.

- DPX to JPEG2000. This is a digital cinema preservation scenario based on the EDCINE project. DPX files are converted into MXF wrapped JPEG2000 with lossless compression to create Master Archive Packages (MAP).

- Temporary archiving of file format video. This is a general preservation scenario whereby if essence is already in a file format, or can be transferred to a file format using a bit level copy (e.g. the DV family), then it should be archived in this format until it is obsolete and then it should be transferred to an uncompressed format.

Overview of video file format migration model

A simple model of video preservation by using migration is shown in Figure 4. The diagram is a model of the following general scenario for long-term preservation of video:

1) An archive has one or more items of video essence (this could be new material that has been broadcast recently or it could be old items on video tape on shelves) that needs to be preserved in the archive. This is the 'START HERE' point at the top of the diagram.

2) The archivist gathers technical metadata on this video, e.g. what format it is in, whether it is on video tape or data tape etc. and assesses whether the format is a problem.

3) If the video is already in a file format, or can be converted without any loss into a file (e.g. DV or IMX D10), and if that format is not problematic, then the video is copied into a file and archived 'as is' on a temporary basis, e.g. for a few years.

4) If the video is in a problematic format, e.g. the format is not file-based or is obsolete, then a new format needs to be chosen. The archivist evaluates the different formats that could be used and assesses the cost of storage for these formats.

5) The archivist then decides to either:
   a. Decode the video to an uncompressed format if they can afford the storage. This is the long-term preservation solution (the END HERE point at the bottom of the diagram).
   b. Encode the video in a compressed, but lossless preservation format (e.g. JPEG2000 lossless). This encoded form can't be considered a long-term solution, so is archived temporarily.
   c. Encode the video in a compressed and lossy preservation format, e.g. high bit rate MPEG2 or MPEG4, because of the prohibitive storage costs of using a lossless or uncompressed format. This encoded form can't be considered a long-term solution, so is archived temporarily.

6) If in step 5 the video has been put into a file in an encoded form that is archived 'temporarily', then the format is assessed again after a few years to see if it is at risk (e.g. because the format is now obsolete). If not at risk then the file it is archived temporarily 'as is' again. If at risk, then a new format is chosen and the process repeated.
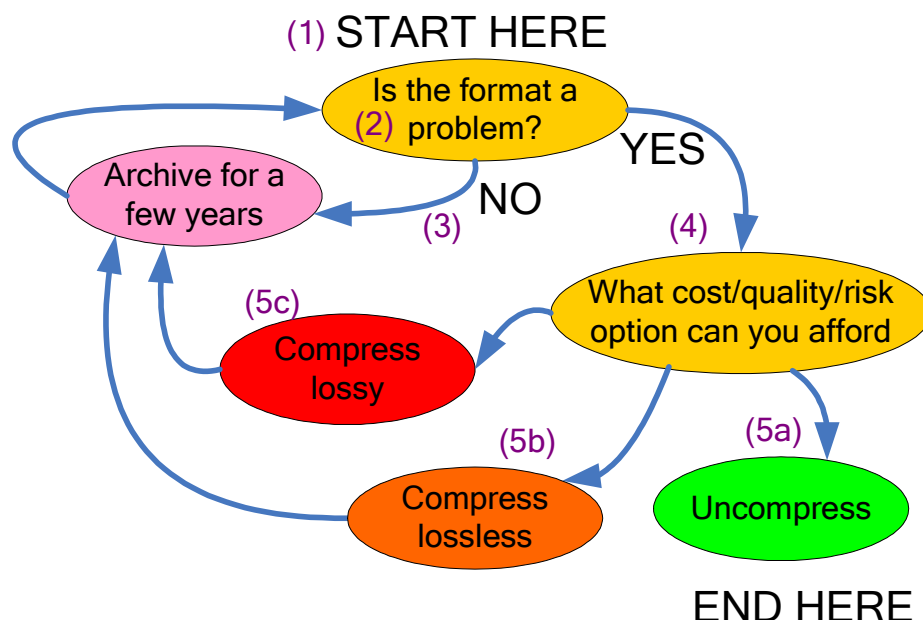

Figure 4 Simplified view of a model of video file format migration

The question of whether a format is problematic is a complex one and involves several factors. It also depends on an organisation's capacity to reduce or mitigate these problems.

A format might be considered problematic if

- The format is considered obsolete from the point of view of users of the content, so, for example end user tools for editing and repurposing audiovisual material may no longer be able to use the content,

- The format is obsolete from the point of view of the archive responsible for maintaining access to the content, so, for example tools aren't available to allow the content to be decoded or transcoded before delivery to end users.

- The format is in a form that is hard to manage from a preservation perspective, e.g. its internal structure makes it difficult to control, manage and check the integrity of the components (metadata, audio, video, timecodes etc.)

- The format requires storage space/performance beyond the capacity of the preservation infrastructure, for example very large files with high bit rates for uncompressed video.

In an ideal world, the format used for preservation should be the highest quality possible and preferably uncompressed. The format should be the one best suited to preservation, e.g. having the longest lifetime, and not dependent on the formats needed for use of the content by end users. However, in reality, uncompressed video is still very expensive to store and there is a cost associated with converting between preservation formats and end user formats on demand, especially for master quality.

This leads to more complex decision making than shown in Figure 4. An example of 'real world' decisions is shown in **Figure 5**. This example shows some of the decision points for RAI. So, for example, an immediate concern with any video format is whether the archive has the tool support to be able to access the content and deliver it to the users of the archive. If tool support is problematic, e.g. because the tool currently used becomes unsupported by a vendor, then there are several options:

1. Keep the format, but switch to another vendor for the tools

2. Keep the format and fix the existing tools (if possible, e.g. they are in-house)

3. Keep the format and develop/maintain new tools (e.g. based on open source)

4. Keep the format and take a multivalent approach (see later for details)

5. Change the format (format migration)

This highlights that a viable preservation strategy is to create and maintain the means to support the tools needed to access and use content, which replaces the need to migrate the format of that content. These tools will typically be some form of decoder that can extract the video signal and related information, e.g. timecodes, metadata, subtitles etc. This approach is particularly viable when the format and tools have either been developed 'in house' and hence full control already exists over the tools, or where there are well developed open source tools that support the format (e.g. ffmpeg) that can be used to kick-start an in-house capability.
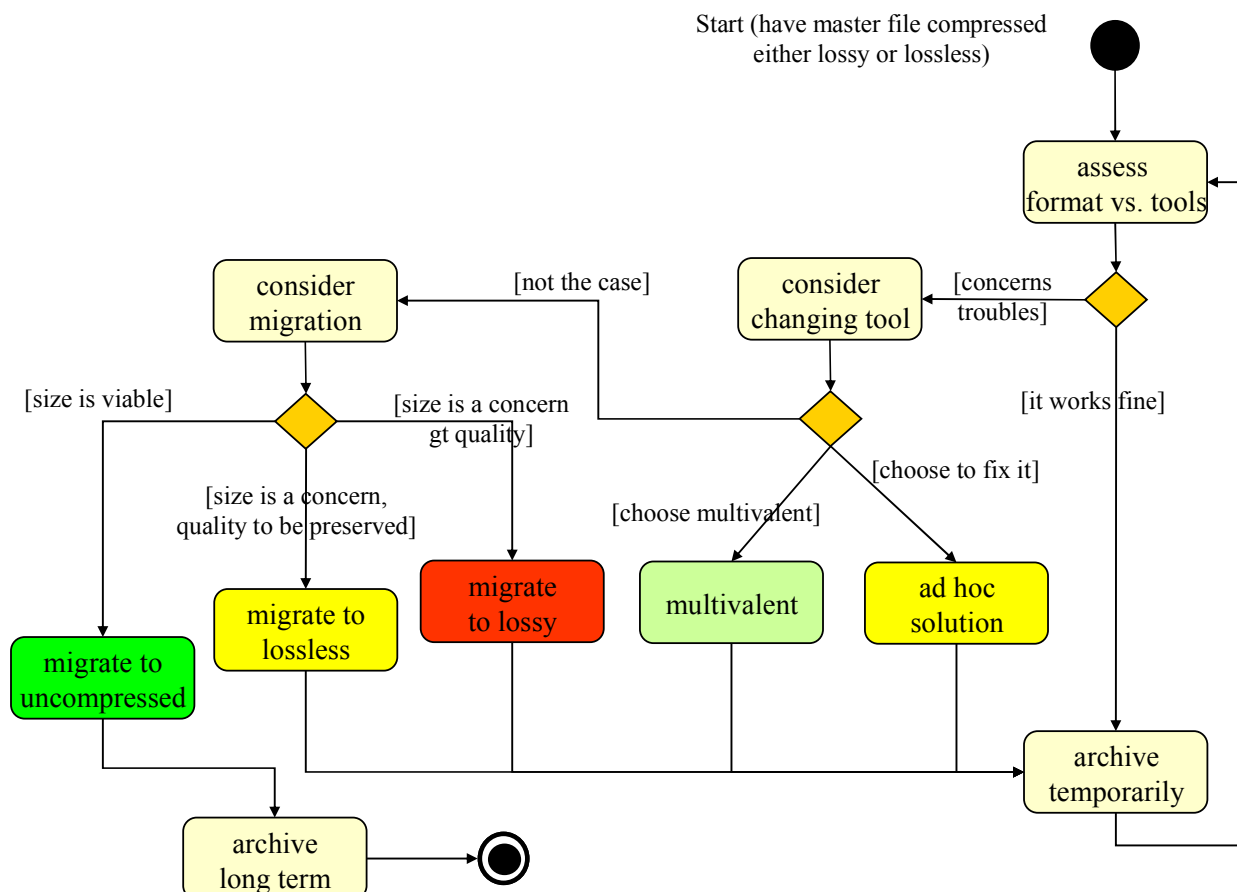
Figure 5 Example of video format migration decisions for RAI

Considerations for assessing format and tool options

The version of the preservation game depicted in Figure 5. applies to the coding schemes of master quality video material and is characterised by a cascade of three decision points dealing with assessment of the coupling of formats to tools, evaluation of options on the tool and evaluation of options on the formats.

Although one route through the diagram is for the preservation process is to end with a migration to the uncompressed domain, it is also conceivable to never get to that point. The goal of a preservation model is to provide the means for evaluating the trade-off points for supporting the required decisions.

The tools referred are the software or hardware components which permit the actual access to master AV content for delivery. The real context of use of such tools may vary considerably depending on the environment (operating system, software dependencies, software and hardware interfaces) and on how much the single components can be isolated and replaced with other ones if necessary.

The simplest tool is a decoder that is able to provide a video uncompressed stream from a coded data stream (does nothing if format is already uncompressed). The decoders must interface with other components such as encoders, video-renderers, multiplexers. The archive must always be responsible for providing the decoding capability of master level, even in the simplified (but temporary) case where the archive format is exactly the same

one required for access/delivery to production and decoding would not be necessary for delivery.

When RAI assess a tool versus a format, then the considerations they apply include:

- The tool works correctly – assuming the archive has N items in format F it is expected that the tool works correctly with all of them. This cannot be taken for granted as the various items might have been ingested from various sources in different times.
- Availability of alternative tools – if other tools are also available for the same format F an occurring problem with a subset of archive items is less critical (otherwise it would be better to reject the item at ingest), however it must be ascertained that all N items in format F are accessible by at least one tool (required). However, if there is only one tool then there is a concern.
- Performance is sufficient – the minimum required level of performance should be given by the SLA, it may vary considerably, and anyhow the performances will also strongly depend on the infrastructure environment. A metric could be defined as the amount (total duration) of material deliverable within a defined time window, assuming a fixed infrastructure. An alternative metric could be the time required for delivering all the items in format F, with the same infrastructure. The latter metric is related to the time necessary to perform a complete migration from format F, and will tend to be critical with the increasing of N (actually the total duration material in format F).
- Authoritative certification - if an independent organisation can be identified as testimonial that the tool works correctly with format F, this can confirm the factor "it works correctly " and lower the need of tool redundancy
- Capability to recover from tool issues – this gives an evaluation of the possibility to recover from a novel problem. Key factors are availability of source code for the complete tool and availability (internally or externally to the organisation) of technical staff able to handle the issue. Alternatively there should be a contract agreement with the relevant technology providers for covering this aspect.
- Criticality versus the environment – a measurement of how much the tool is depending on a given environment. E.g. it is more critical if it is bounded to a specific version of a specific operating system.

The ideal preservation model defines a "format vs. tool" assessment which permits a "no loss scenario". The assessment cannot be done in isolation from the environment either: it needs to be performed within a given time window and but also in the occurrence of events which could change the options, for example:

- Planned modification of environment/infrastructure
- Modification of parameters on performance requirements
- Rate changing from certification authorities
- Change status of availability of alternative tools
- Increase of rejections at ingest of items for validation against format F

In case of negative result of tool assessment against format, the flow chart of **Figure 5** proposes to consider firstly an action on the tool and secondly a format migration. Actually this is only for a greater clarity of the picture, because there should be a single decision point for selecting the way forward.

The decision is shown as taken on a "costs vs. content quality" criterion, since these are typically the dominant factors. However, other factors are known to require consideration:

- Time – how much time will be necessary for the solution to be effective and which are the consequences in the meantime
- Success probability – does the option under consideration give certitude of success or there is risk of failure to be considered. This factor is related to the cost.

Video file format migration model

A model of video preservation by using migration is shown in Figure 6. The diagram uses UML notation. The rounded rectangles are activities, i.e. things that are performed. Activities are connected by arrows that show the order in which the activities are undertaken. The diamonds are decision points, i.e. where there are two or more options to choose from. Each arrow that comes out of the diagram is labelled with the decision/option for that path. The diagram is a model of the following general scenario for long-term preservation of video:

1) An archive has one or more items of video essence (this could be new material that has been broadcast recently or it could be old items on video tape on shelves) that needs to be preserved in the archive. This is the 'start' state.

2) The archivist gathers technical metadata on this video, e.g. what format it is in, whether it is on video tape or data tape etc.

3) If the video is already in a file format, or can be converted without any loss into a file (e.g. DV or IMX D10) then the video is copied into a file. If the file is in a video tape format (e.g. D3 or DigiBeta) where there is no direct access to the video at the bit level, then the video is decoded and put in a file in a raw format (e.g. recording the SDI output of a video player into a file).

4) If in step 3 the video has been put into a file in an encoded form (e.g. copying the MPEG2 into a file from an IMX tape), then the format is assessed to see if it is a problem (e.g. because the user community can't use the format).

   a. If the format itself is not a problem, the maintaining tool support might be, so this is assessed next. If tool support is a problem, then the archive assesses whether to fix the tools, replace with tools from a different source, or develop their own set of tools.
   b. If the archive is confident that neither the format nor the tools are an issue, then the file it is archived temporarily 'as is'. Note: 'temporarily' could mean 5-10 years, or more.

5) If in step 3 the video has been uncompressed to a file or it is in a format that is a problem, then the archive needs to decide how best to store it, e.g. to change format - they may not be able to afford to keep it uncompressed, or they may want to change the encoding to something that is better for preservation (e.g. JPEG2000 lossless).

6)   If the archive decides they should change format and use compression (lossy or lossless) then the video is transcoded and stored in the archive temporarily in this format pending a further migration in the future, ideally to uncompressed.

7)   If the archive decides they should keep the video uncompressed then no further format migrations are necessary and the process stops here (not strictly true as everything becomes obsolete eventually, even uncompressed video). This is the 'finish' state of the diagram.
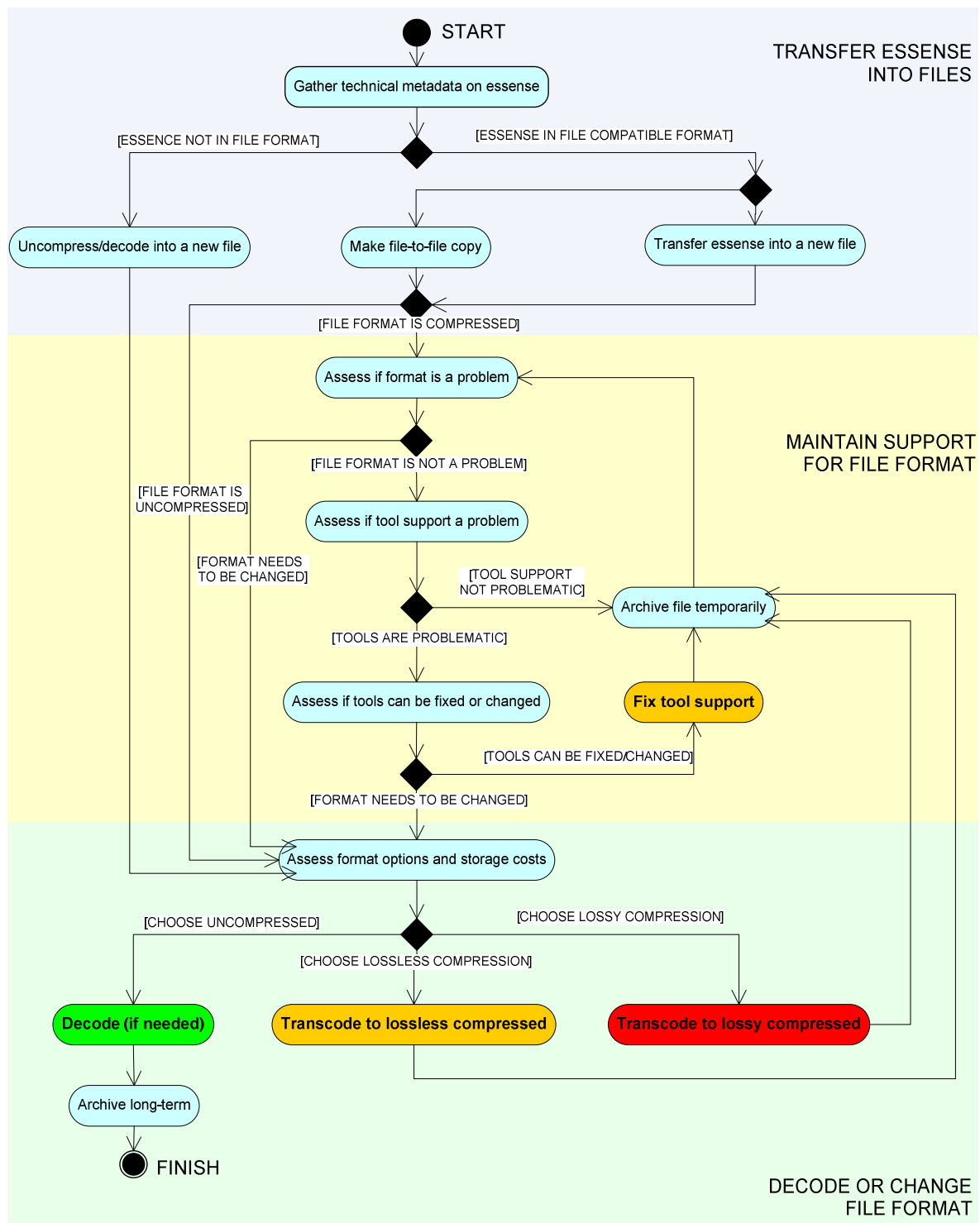
Figure 6 General purpose video file format migration model

Figure 6 is divided into three zones.

1. The zone at the top (light blue) is concerned with transfer of video from its existing format into a digital file format, for example transferring content from IMX video tapes into MXF files stored on data tape or hard drives.

2. The zone in the middle (pale yellow) is concerned with keeping files in the current format for as long as sensible. This is where 'temporary archiving' takes place. The activities in this zone are on going assessment of the file format and tool support and any actions that may be needed, e.g. changing/fixing/developing tools to allow the content in the files to be accessed.

3. The zone at the bottom (light green) is concerned with changes in file format, e.g. decompressing content, or transcoding content (lossy or lossless). The decision on what format to choose is done in conjunction with storage options. For example a decision might be made that it is now affordable to move to uncompressed.

The main preservation approaches used are colour coded in green, orange and red. These strategies are:

(a) maintain tool support for the current format

(b) migrate file formats (transcoding)

(c) use uncompressed (decoding)

Green is used to indicate a 'safe' strategy, i.e. decoding is recognised as a safe long-term option as the content is in a very simple form (making use or recoding very easy), it is relatively robust to problems with storage (e.g. bit rot), and the standards used for uncompressed content (e.g. as used within the SDI specification) are relatively long-lived.

Orange is used to indicate a strategy where there are some potential risks or concerns, e.g. the formats/tools chosen may have a limited life, or the decision to fix/maintain tools could be costly or error prone.

Red is used to indicate a strategy where there are potentially serious concerns (it can still be a viable strategy nonetheless), for example lossy transcoding will introduce loss of quality by definition and even if this seems to be imperceptible now it could turn out to be a very real problem in the future or when further transcoding is done with codecs that don't 'work well together'.

Clearly, a sensible approach is to avoid lossy transcoding if at all possible.

Possibly the simplest and lowest risk strategy is for the initial transfer of content into files to be done 'as is' without changing the video encoding and then keep this format until it is cost effective to decode the video into an uncompressed form for long term archiving.

The reality may well be somewhere in between these two extremes and involve efforts to maintain tools so original formats can be used as long as possible, or for transcoding to be done to lossless compressed formats (e.g. JPEG2000) to minimise storage costs.

2.2.2.   D3 video preservation example

A simple example of video preservation in files is the D3 at the BBC as shown in Figure 7. D3 is a video tape format that can't be transferred directly to a file. Therefore, it is ingest by playing the tape with an SDI output and then recording the video into a file. The BBC decides to store the video uncompressed so no further format migrations are needed
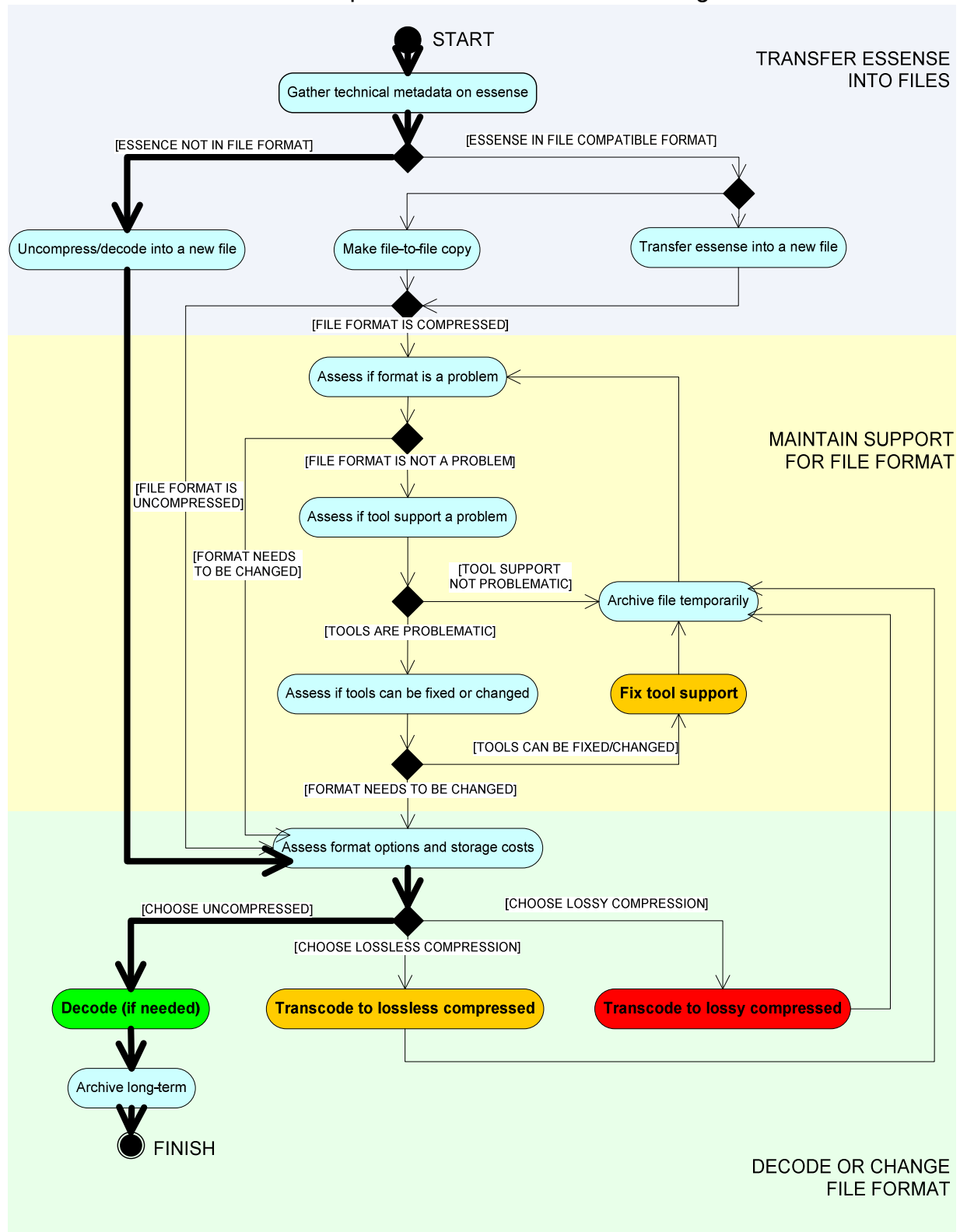

Figure 7 Example of preservation of D3 video at the BBC

2.2.3.   *IMX video preservation example*

Figure 8 shows a possible approach for IMX video tapes (based on a RAI scenario). The first step (1) is to transfer the video from the IMX tape into a file. This is a bit-level copy and doesn't involve decoding and then encoding again. This file (after wrapping in MXF) contains compressed video (MPEG2 using the D10 standard), but is not considered problematic (2) so is archived temporarily without modification (3). This is the best strategy as it involves no changes at a bit level to the content that was originally on the IMX tape. However, at some time in the future (say 10 years), MPEG2 will be considered obsolete, which means the format needs to be reassessed (4) and a format migration is needed. However, due to rapidly falling storage costs, in ten years time the archive is able to afford to store as uncompressed (5), so the MPEG2 files can be decoded (6).
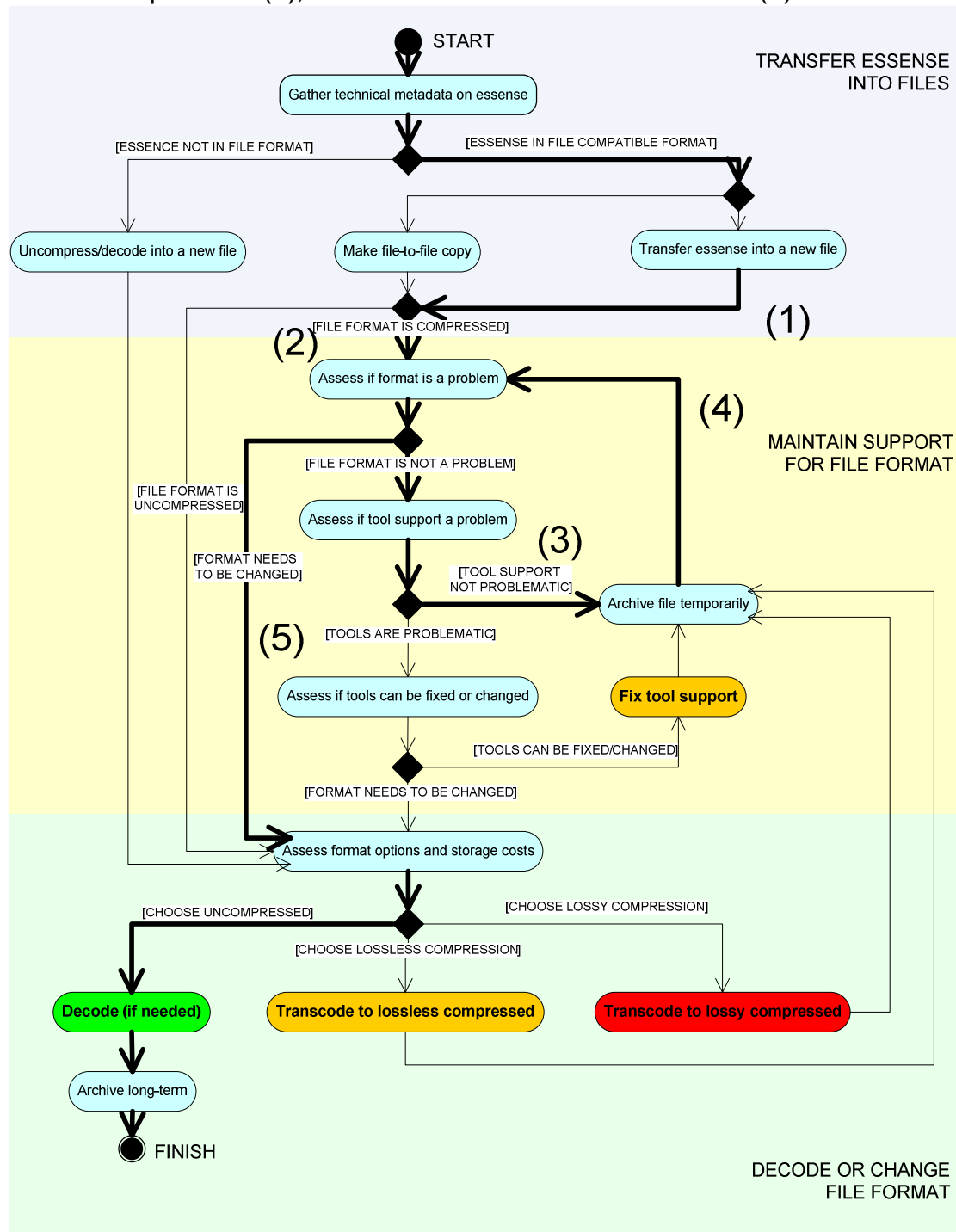

Figure 8 IMX video preservation example at RAI

2.2.4.   *JPEG2000 digital cinema video preservation example*

Figure 9 shows the model applied to a Digital Cinema scenario using JPEG2000 lossless as a preservation format. The scenario starts with DPX files which are copied into the archive directly (1,2), but are considered inappropriate as a preservation format (e.g. too expensive to store uncompressed) so need to be encoded. JPEG2000 lossless is the initial choice (3) and these are archived temporarily, but will become obsolete eventually (4,5) but by that time the archive will be able to afford to store the video uncompressed (6).
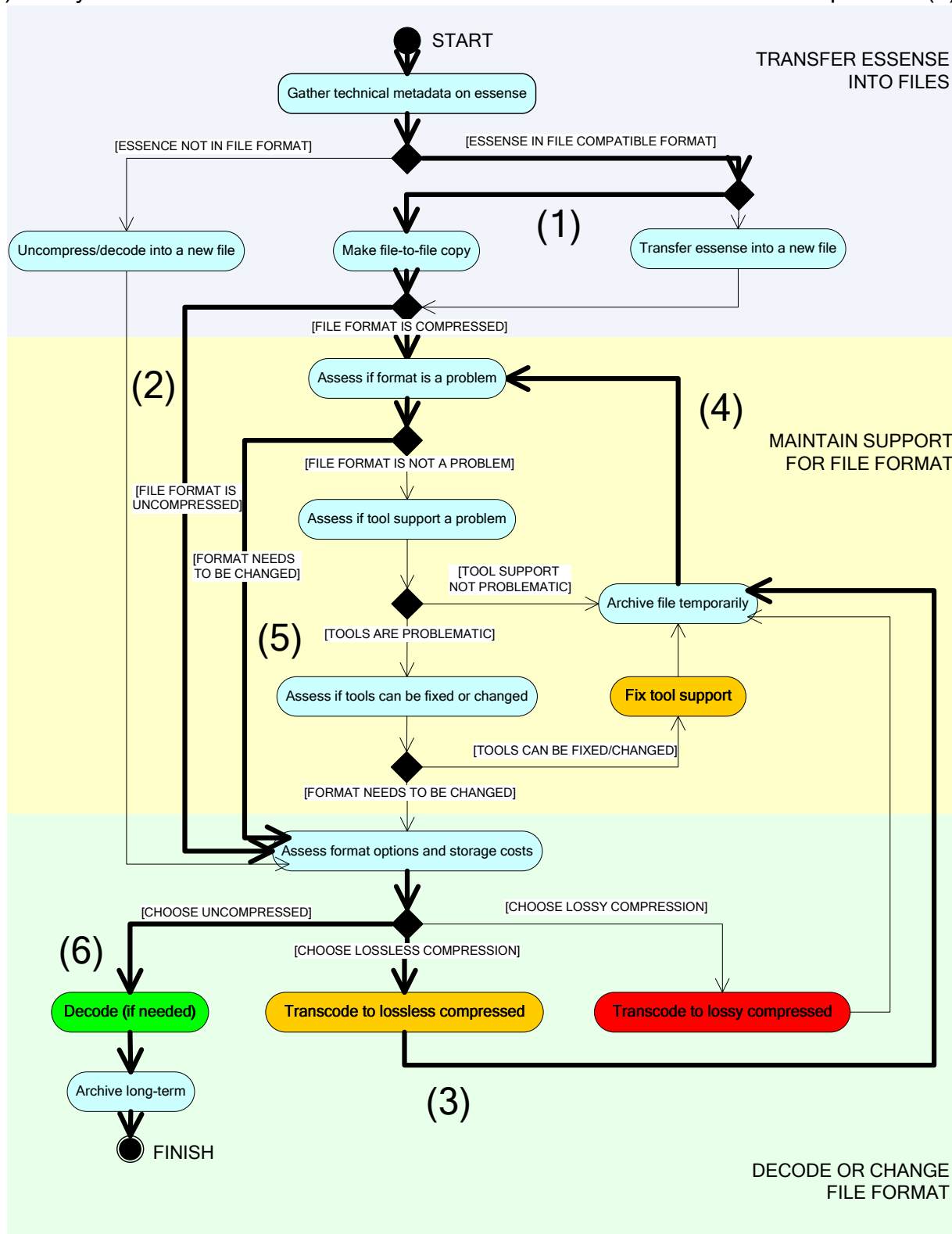
Figure 9 JPEG2000 Digital Cinema video file format preservation example

2.2.5.   U-Matic video preservation example

Figure 10 shows an example of how lossy transcoding can have a role in preservation of video. Suppose an archive is digitising old analogue video that isn't of particularly high quality (e.g. U-Matic, VHS). Using a lossy compression format can make sense (e.g MPEG2 at 50MBit/sec) as a lot of storage space can be saved with very little qualitative loss of quality (higher bit rates or lossless compression only serve to capture the noise in the original signal). Therefore, the first stage is to digitise the video (1) and then compress to a lossy file format (2) that is stored temporarily (3). Eventually, there will be a need to migrate from this format (4) and at this point, along with the other scenarios above, moving to uncompressed is likely to be affordable (5,6).
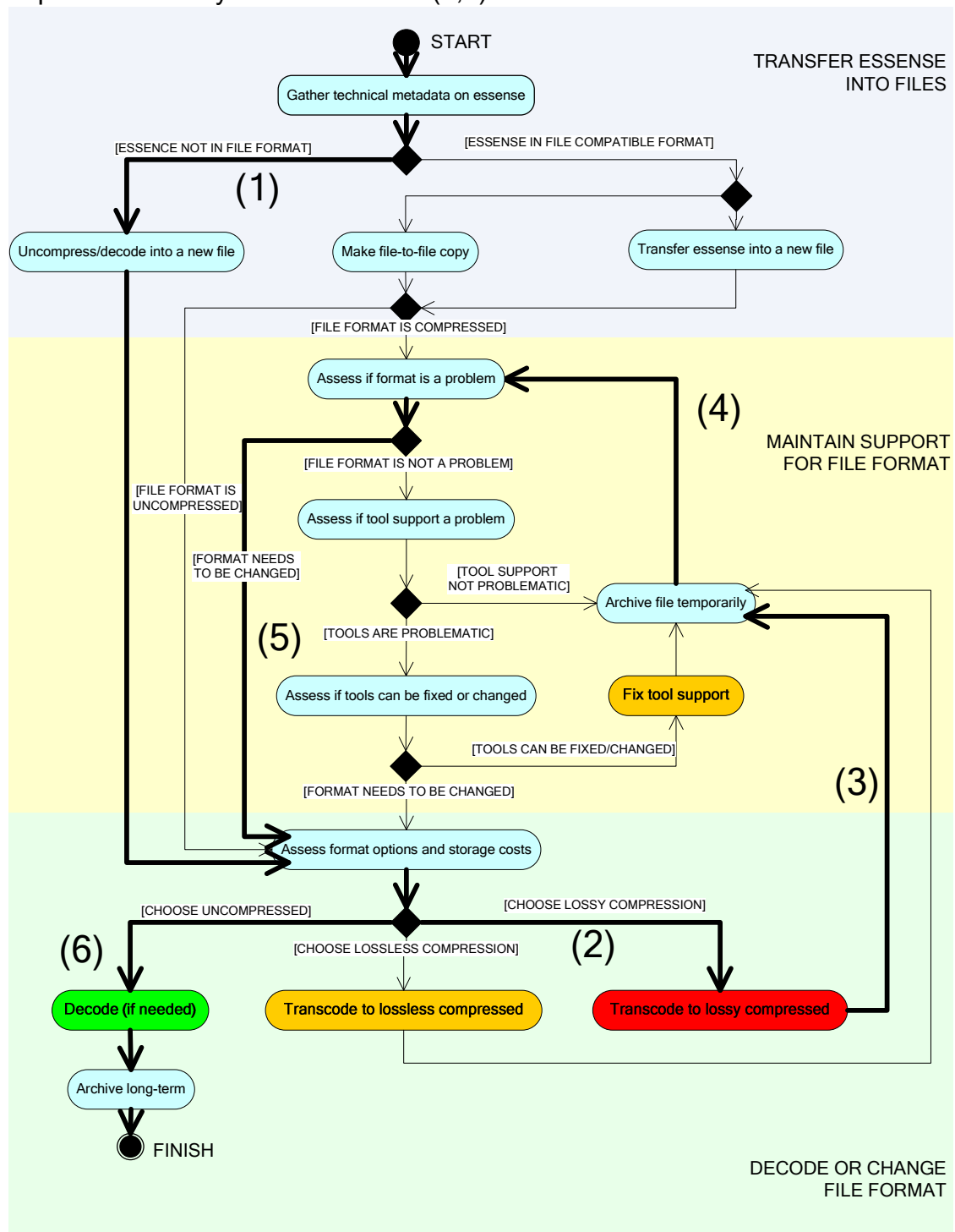

Figure 10 U-Matic video preservation example

2.2.6.   Bit level copy of content from Digital Video (DV) tapes

In the preceding sections we asserted that DV55 is an example of a video format where it is possible to transfer the content from digital video tape into a file in a bit perfect way, i.e. the bits in the file are identical to the bits as recorded on to the tape.

Technically this is possible by using FireWire[56] transfer from the video device, e.g. camera to a system that then records the DV stream into an appropriate wrapper, e.g. an AVI file. However, it should be noted that depending on the specific equipment used to play the tape, the bits as delivered by the equipment may not be the same as the bits as originally recorded. Error correction technology in the device may automatically detect and conceal errors in the bit stream. This is used, especially for the more consumer oriented formats (miniDV) due to the somewhat error prone nature of the media, e.g. because of the very high recording density.

The hardware used to play the DV content and hence make the 'bits' accessible through firewire may or may not apply correction and the type of correction can vary. It may not be possible to determine what correction has been applied until after the transfer has taken place. Thankfully tools exist to analyse DV files and hence assess to what extent the bits have been adjusted in order to conceal errors in reading the tape within the device. DV Analyzer[57] is a freely available example.

DV Analyzer will report on whether video error concealment has taken place. Types of concealment reported include a STA 'STAtus of compressed macro block' value:

- STA = 'A': This block of video data is replaced by the corresponding video data of the previous frame (very common)
- STA = 'C': This block of video data is replaced by the corresponding video data of the next frame
- STA = 'E': Unspecified concealment
- STA = 'F': Error. This typically is represented as grey blocks on the frame.

Therefore, whilst it is theoretically possible to do a bit level direct copy, in practice care must be taken to ensure this is actually achieved, and if not to understand and record exactly what concealment has taken place.

However, this also raises further questions on what makes the 'best' transfer.

- If just the raw bits as originally recorded onto the tape are transferred to a file then artefacts may be present due to errors when the tape was read. The ability to correct these errors at a later date, e.g. using software, may be limited, especially if the device supplying the data does not indicate where the errors exist.

- If the device does do error concealment, then exactly what has done may not be known (e.g. STA type E above) and may prevent a better error correction approach being applied in the future. On the other hand, because the concealment is done at time of transfer within the device and at the time that errors occur, then the

---

[55] http://en.wikipedia.org/wiki/DV
[56] http://en.wikipedia.org/wiki/IEEE_1394_interface
[57] http://www.avpreserve.com/dvanalyzer

concealment may be better than what could be subsequently be achieved using external systems.

- Professional DV devices, e.g. DVCPro will typically have an SDI output. This provides an alternative way of capturing the content of the DV tape. In this case, further video processing may be applied by the device in creating the SDI signal. This form of transfer should not be considered as a 'bit level' transfer in any way, but it does result in a digital representation in a file (e.g. see D3 example above). Again, this may be better or worse than corrections that might be applied at some point in the future to a bit level transfer using FireWire.

The 'best strategy' to use is under debate[58]. Some of the issues go back to the fundamentals of 'what is preservation', e.g. is it better to capture the video (SDI) output of the device as this is as close as you can get to the way the video would have been played back and hence seen at the time, or is it better to strive for a bit level copy of the recorded data, even if that does contain errors that aren't in any way concealed, as this is as close as you can get to the 'raw data'. We don't make a recommendation on this here. A good discussion of the issues is provided by Audiovisual Preservation Solutions in their paper "Digital Tape Preservation Strategy: Preserving Data or Video?"[59]

The only recommendation that we make is to take great care in verifying any assumptions that a 'bit perfect' copy of data from a digital video tape really is a perfect copy and to be sure that this is actually the best thing to do.

2.2.7.   File Formats: Risk of Loss and Sustainability Factors

It is common in the preservation community to speak of technical obsolescence or risk of loss, e.g. a particular format no long has players readily accessible in the market place and hence there is a risk of not being able to access content in that format. Risk of loss can come from many sources, including obsolescence, e.g. not having expertise, costs being too high, a format being particularly sensitive to media degradation etc... Migrating to a new format is one way of reducing the risk of loss associated with a 'legacy' format.

Stanford's video identification guide[60] is a good example of categorising according to risk. This mainly applies to discrete analogue media. The Library of Congress consider a range of sustainability factors[61] for file formats, which are in effect the inverse of risks in that they consider what is necessary for there not to be a risk of loss. Factors include:

- **Disclosure**, which refers to the degree to which complete specifications and tools for validating technical integrity exist and are accessible to those creating and sustaining digital content.

---

[58] There is a long thread on this topic on the AMIA list with excellent points made by Jim Lindner, Richard Wright, Park Seward, Isaiah Beard and others. See the thread etitled "[AMIA-L] Best practices for Mini DV preservation and digital transfer."
[59] http://www.avpreserve.com/dvanalyzer/dv-preservation-data-or-video/
[60] http://videopreservation.stanford.edu/vid_id/index.html
[61] http://www.digitalpreservation.gov/formats/sustain/sustain.shtml

- **Adoption**, which refers to the degree to which the format is already used by the primary creators, disseminators, or users of information resources.

- **Transparency**, which refers to the degree to which the digital representation is open to direct analysis with basic tools, including human readability using a text-only editor.

- **Self-documentation**. Digital objects that are self-documenting are likely to be easier to sustain over the long term and less vulnerable to catastrophe than data objects that are stored separately from all the metadata needed to render the data as usable information or understand its context.

- **External dependencies**, which refers to the degree to which a particular format depends on particular hardware, operating system, or software for rendering or use and the predicted complexity of dealing with those dependencies in future technical environments.

- **Impact of patents**. Patents related to a digital format may inhibit the ability of archival institutions to sustain content in that format. Although the costs for licenses to decode current formats are often low or nil, the existence of patents may slow the development of open source encoders and decoders and prices for commercial software for transcoding content in obsolescent formats may incorporate high license fees.

- **Technical protection mechanisms**. To preserve digital content and provide service to users and designated communities decades hence, custodians must be able to replicate the content on new media, migrate and normalize it in the face of changing technology, and disseminate it to users at a resolution consistent with network bandwidth constraints.

The LoC have applied these to a wide range of moving image formats[62] This includes both wrappers (e.g. QuickTime and MXF) and video/audio encoding. For example, MXF-wrapped JPEG2000 is considered by the LoC to be a relatively sustainable formats as the standards are open, the wrapper is transparent, adoption is already good or is growing, there are freely available implementations etc…

On the other hand, consider Apple's QuickTime ProRes format[63], which is widely used in the production and post-production community, in particular in conjunction with the Final Cut Pro[64] tool, but not currently included on the LoC site. QT as a wrapper is not an issue as this is widely used and the specification is open. ProRes as a codec satisfies the 'adoption' criteria and there is reasonably good tool support, including free decoders for Windows and Mac. The specification is proprietary, which means the 'disclosure' criterion is only partly addressed. However, the other criteria of patents, external dependencies, self-documentation and transparency all mean that Apple ProRes cannot be considered a sustainable preservation format.

---

[62] http://www.digitalpreservation.gov/formats/fdd/video_fdd.shtml
[63] http://images.apple.com/finalcutstudio/resources/white_papers/L342568A_ProRes_WP.pdf
[64] http://www.apple.com/finalcutstudio/finalcutpro/

The issue of patents is also an interesting one and the LoC note that "The core components of emerging ISO formats such as JPEG2000 and MPEG4 are associated with "pools" that offer licensing on behalf of a number of patent-holders. The license pools simplify licensing and reduce the likelihood that one patent associated with a format will be exploited more aggressively than others. However, there is a possibility that new patents will be added to a pool as the format specifications are extended, presenting the risk that the pool will continue far longer than the 20-year life of any particular patent it contains. Mitigating such risks is the fact that patents require a level of disclosure that should facilitate the development of tools once the relevant patents have expired."

Beyond the issues of sustainability, there is also the question of when it makes sense to change file format from a business perspective. For example, several archives in PrestoPrime use MPEG1 video (ISO/IEC-11172-2) as a browse quality file format. The standard is already 17 years old, it's currently not recommended in any particular application, but decoding is however fully supported in a wide range of software, furthermore the standard includes reference software which gives the possibility to re-build an emergency decoder. Therefore, whilst it could be considered 'obsolete' from the point of view of no one using it anymore, existing content in MPEG1 format isn't at risk. However, from a user perspective, Macromedia Flash or MPEG4 are more 'useable' and hence archives are now transcoding their MPEG1 to more contemporary delivery formats or going back to the masters to generate new browse copies. Whilst fine for browse resolution, this does highlight the need to consider the user community as well as preservation needs. Having to transcode a preservation format in order to deliver an access format has a cost associated with it that has to be considered in the overall preservation and access strategy of the organisation.

As a final point to throw into the mix of 'when to migrate between file formats', it is worth noting that encoding/decoding can get 'better' with time (see Annex II for an example). For example, as the compute power available increases from the platforms that the encoders/decoders run on, then more 'number crunching' can be done for the same cost (or within the same time). Increasing the quality of encoding (e.g. Multi-pass Variable Bit Rate encoding[65]) can become viable. This is also one reason why new codecs emerge in the first place, for example, MPEG4 is computationally much more expensive than MPEG2 and it is only relatively recently that MPEG4 is viable for high bit rate material. Likewise improvements may be made to the encoders themselves for a given format. Decoding quality can also vary between different implementations of the same specification and as applied to the same file. In particular, the more knowledge the implementer of the decoder has of how exactly the encoding has been done, the better the decoding can be done. Not all that knowledge can come from examination of an encoded file and hence it is always worth considering whether it is possible to capture and archive the encoder in a transparent form at the same time as archiving the content that has been encoded with it. Archives are already seeing some of these issues. For example in a mass migration of MPEG1 browse quality content to MPEG4, INA have seen that early MPEG1 files weren't of a sufficient quality for direct transcoding to MPEG4 and hence the master copy had to be used, but for later MPEG1 files a direct transcode was of acceptable quality. This shows how the quality of encoding varied during the period it took to create the original MPEG1 files.

---

[65] http://en.wikipedia.org/wiki/Variable_bitrate#Multi-pass_encoding_and_single-pass_encoding

Archives are therefore faced with an interesting dichotomy. If they decide to migrate a file format too late and there is the risk that suitable decoders/transcoders are not available or can't be easily used. Decide to migrate too soon and there is the risk that decode/transcode quality is not as high as it theoretically could be. The longer the decision is held off, the cheaper storage and processing power gets. Cheaper processing potentially means higher quality transcoding. Cheaper storage means that transcoding may not be necessary at all and the content can simply be uncompressed. Somewhere in between 'too soon' and 'too late' is a 'window of opportunity' where file format migration is most cost effective. Future PrestoPrime outputs will investigate this window further.

Not withstanding all the issues of when to migrate, most archives have a simple question 'what should I migrate to?'

The answer is fairly clear:

- If the content is already in digital file format and is sustainable (see above) then keep it unchanged.

- If you have to change format and can afford uncompressed, then this is the first choice.

- If you have to change format and can't afford uncompressed, then lossless compressed using an open format, e.g. JPEG2000 is the next best thing.

- If you have to change format and can't afford lossless compression, then choose a lossy format that is long-lived and open (see LoC website) and has a quality that matches the source material.

To illustrate the last point, some example migration formats are shown in Table 2. This table comes from the PrestoSpace preservation wiki[66] where further information can be found.

Ingest Format	Migration format	Notes
Low quality media		
VHS tape	DVD	**Access** Perfectly adequate for VHS playback
VHS tape	MPEG-4 files	**Access** Adequate for quality. Minimum data rates (MPEG-4): 500k b/s. There are MANY potential access formats, and they come and go.
VHS tape	DV files	**Archive** (temporary) 25 M b/s, 12 GB/hr. Migrate to lossless for preservation.
'low end' digital files	Save as is, AND save as DV or lossless	**Archive** (temporary) Before format or DV format becomes obsolete, migrate to lossless for preservation.
DVD	DV files	**Archive** (temporary) 25 M b/s, 12 GB/hr. Migrate to lossless for preservation.

---

[66] http://wiki.prestospace.org/pmwiki.php?n=Main.Roadmap

Medium quality		
U-Matic	DVD	**Access** Reduces quality; suitable only for viewing
U-Matic	DV files	**Archive** (temporary) 25 M b/s, 12 GB/hr. Migrate to lossless for preservation.
DV, DVCAM	DV files (meaning .avi files with native coding)	**Archive** (temporary) transfers to computer at 25 M b/s, resulting in an .avi file 'clone' of the original DV tape. Migrate to lossless for preservation.
High Quality		
BetaSP, Digibeta, other pro formats	Uncompressed	**Archive** Uncompressed standard definition video: 200 Mb/s. About 100 GB for one hour, meaning 25 DVD-ROMs (or part of one data tape or hard drive).
BetaSP, Digibeta, other pro formats	Motion JPEG 2000 (lossless version)	**Archive** lossless compression, with a resultant data rate of around 90M b/s. About 40 GB for one hour.
DVCPRO50	.avi files, DV coding	**Archive** (temporary) As for DV, but at twice the data rate. Less susceptible to loss on future migrations. Migrate to lossless for preservation.
'High end' digital files	Save as is	**Archive** (temporary) Before format becomes obsolete, migrate to lossless for preservation.

Table 2 Format migration roadmap

2.3. *Preservation using Wrapper Migration*

Initial text from RAI The diagram in Figure 11 shows that a simple process can be defined for wrapping formats (e.g. MXF, QuickTime, AVI). The transwrapping process, by definition, only changes the wrapper and not the encoding of the essence within the wrapper. This preserves the originally essence quality and, substantially, doesn't affect the required storage capacity. Therefore, the considerations to make and the choices to make are the suitability of the tools rather than issues of storage costs.

In many cases a tool may provide at the same time dewrapping and decoding functionalities. In case of raised problems with a given format F, it will be very important to determine which functionality is affected and which is the actual reason of the potential problem. The issue of information (metadata) loss risk in transwrapping processes is not discussed here in details, assuming that such information would be extracted and preserved apart, although it might be beneficial to have it embedded in the AV-media file.
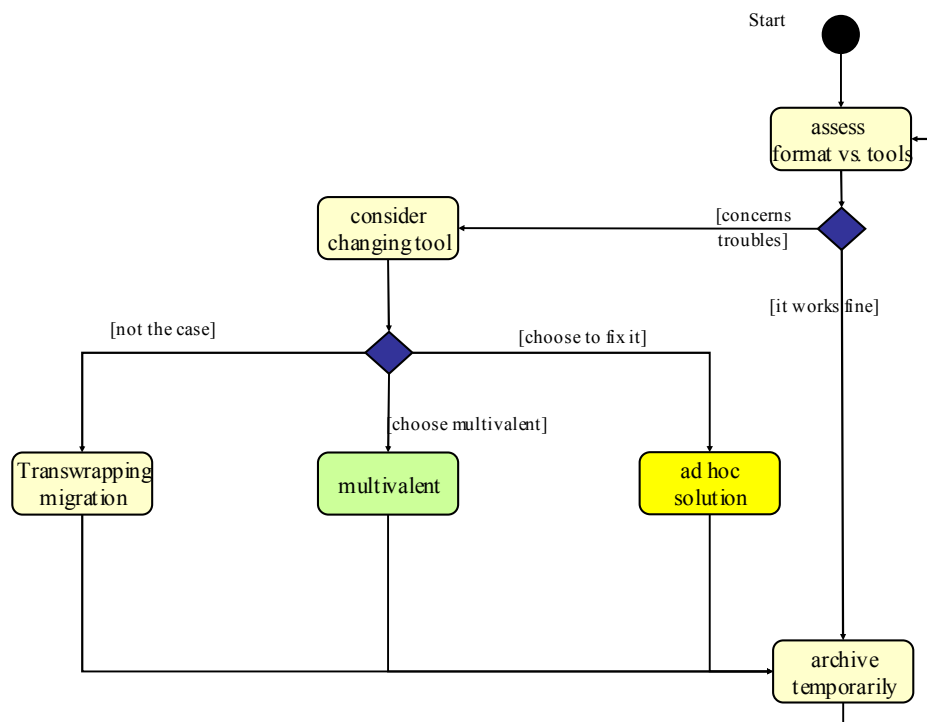
Figure 11 Wrapper preservation decision making process

2.4. Preservation using Multivalent

2.4.1. Overview

In audiovisual preservation when faced with the challenge of maintaining access to audiovisual essence over time, there are two main approaches:

1. Maintain a decoder that can 'understand' the file format of the essence and extract the content when required and deliver it in a form that can be used.

2. Migrate between file formats when there is a risk that the ability to decode the files might be lost or when a different format is required for other reasons (e.g. more efficient storage)

Maintaining a decoder, so for example it continues to work correctly on newer infrastructure (operating systems, storage systems etc.) can require considerable expertise. This expertise may be specific to the particular file format that the decoder is used against. People with the required skills can often be specialists in their field and relatively few in number. If this expertise becomes unavailable (e.g. if an archive, manufacturer or community decides to no longer use or support a codec) and consequently tools to decode essence in a particular format become obsolete, then it can be costly and error prone to try to develop or source this expertise in order to continue to use content in a particular format.

Multivalent provides an alternative approach.

Put simply, a decoder is written for a particular file format in a way such that the decoder can run on a virtual machine. The virtual machine (e.g. a Java Virtual Machine) abstracts the decoder from the details of the underlying operating system and storage system. If these underlying layers change, then the virtual machine is updated and no changes are needed to the decoder. The decoder is archived along with the files. Every time the file needs to be accessed, it is decoded by the using the decoder running on the virtual machine.

In this approach, there is no need to change (e.g. migrate or perform maintenance) on either the file format or the decoder.

In a sense, the use of a virtual machine allows the decoder to become very long-lived (provided of course that the virtual machine itself is maintained).

The problem is now one of maintaining the virtual machine, i.e. this is the thing that is updated or migrated. Migrating or maintaining a virtual machine is potentially a less costly and less risky approach for the following reasons.

- Unlike the file format or its decoder, which are both by their very nature AV specific, the virtual machine is general purpose and doesn't necessarily require such specialist skills in order to keep it functioning.

- If a virtual machine such (e.g. a JVM) is used that also has a very wide range of other applications, then there will be a very large community with a vested interest in its upkeep. This lowers the risk of virtual machine obsolescence and also allows community based approaches to its ongoing development.

- One virtual machine can support many different decoders. Therefore, in a mixed archive with many file formats, rather than have to maintain a set of decoders, one for each format, there is only the need to maintain one virtual machine.

So, in effect, the problem of needing to maintain a large number of very specific AV decoders and file formats has been swapped for a potentially simpler and less costly or risky problem of needing to maintain just one virtual machine. This is shown diagrammatically in Figure 12.

This is not to say that there aren't risks with this approach, or other costs that need to be considered; only that multivalent provides an interesting and complementary approach to the more traditional ones of file format migration or maintaining decoders.
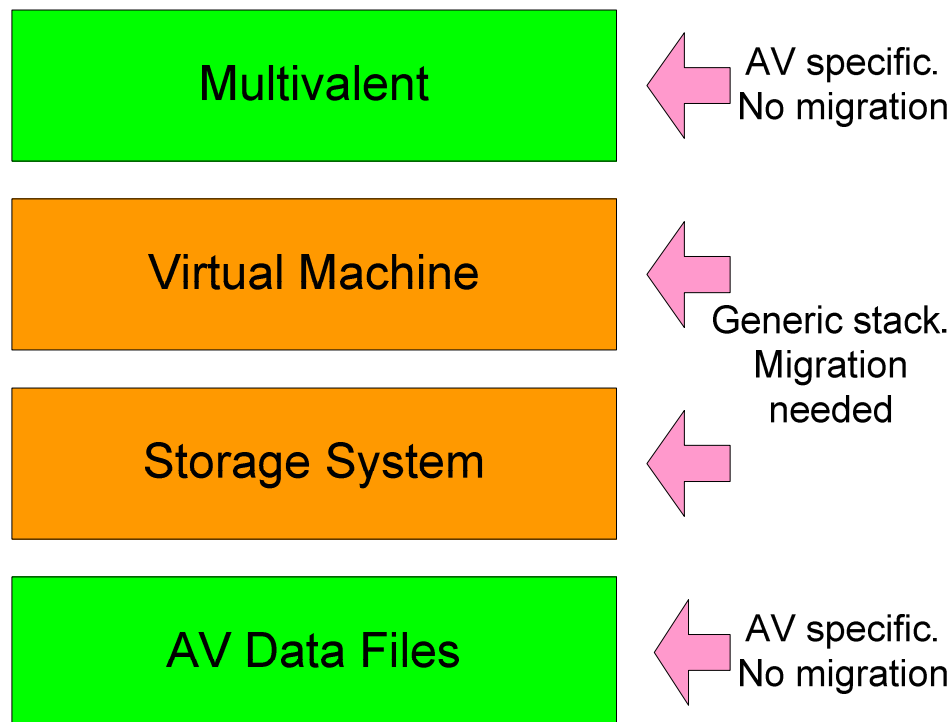
Multivalent ← AV specific. No migration

Virtual Machine ←

Storage System ← Generic stack. Migration needed

AV Data Files ← AV specific. No migration

Figure 12 Diagram showing main elements of the multivalent 'stack'.

In Figure 11, the blocks in green show the AV specific components that do not undergo any migration. The orange blocks show the generic IT components, e.g. these might be a JVM and a HDD server, which will need migration over time, but do not need specific AV knowledge in order to do so.

2.4.2. Multivalent in the context of AV file format migration

In the context of the simple model of preservation using video file format migration shown in Figure 4, multivalent provides an alternative to using transcoding, either lossy or lossless. Instead of transcoding video into a new format, it can be left unchanged and instead a multivalent approach used to maintain the ability to run the decoder for this format. This is shown in Figure 13. Multivalent is shown with green colour coding as it is (theoretically) better to use this approach than transcoding the video format. This is because no losses or transformations occur to the video content. This isn't to say that there aren't risks with multivalent, only that they are purported to be lower than migration
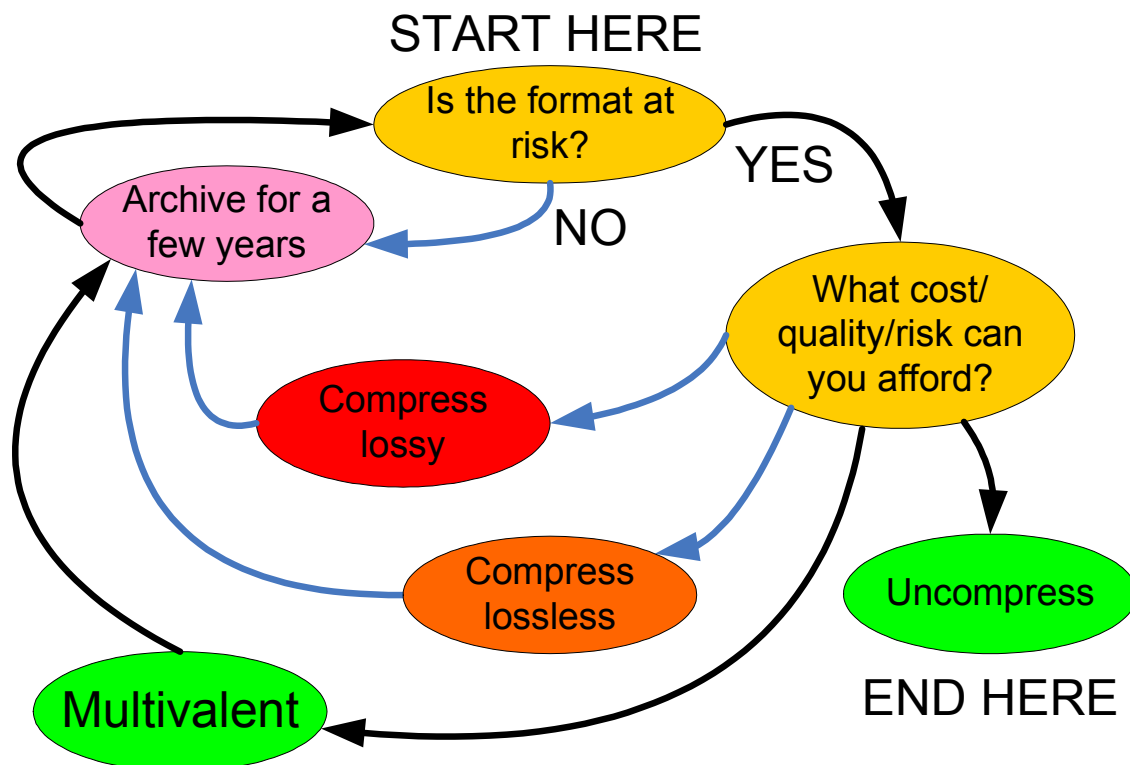
START HERE

Figure 13 Multivalent as an alternative to file format migration

Multivalent shouldn't be viewed as a long-term solution – it only needs to be used for as long as it takes for uncompressed storage of the video to become viable. At that point, there is little need to keep using multivalent for the obsolete format due to the extra costs and risks of maintaining the multivalent system.

In this way, multivalent fits into the model as an alternative to either lossless transcoding or lossy transcoding. When using multivalent to avoid either lossy or lossless transcoding then it can be seen that the video will only exist in two forms: (a) the initial file format and (b) uncompressed.

An expanded view of multivalent in the context of both file format migration and maintaining decoders in the traditional way is shown in Figure 14.
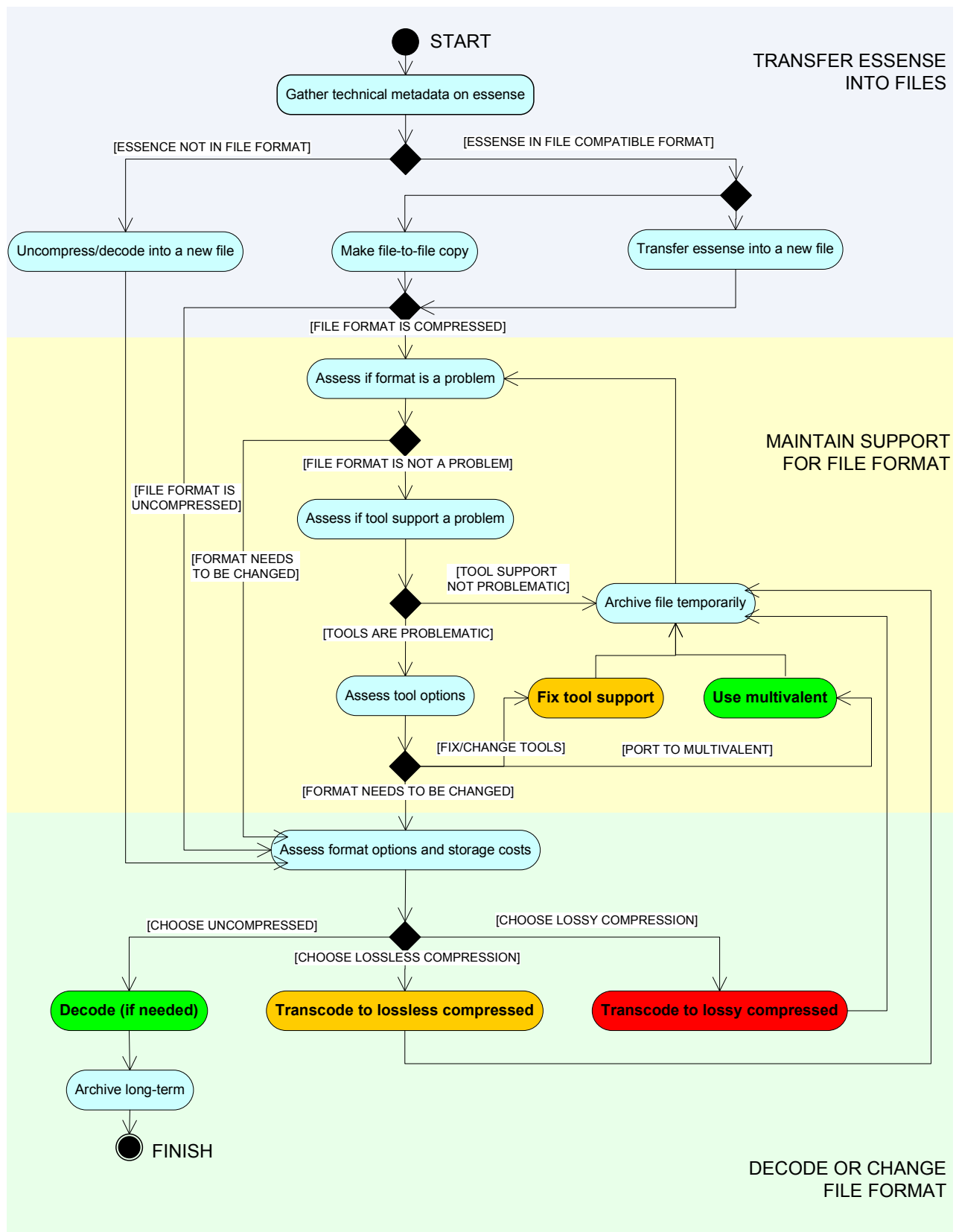
Figure 14 Multivalent in the context of other strategies for preserving video in file format

2.4.3.   *Multivalent video preservation example*

With reference to the example of IMX preservation in Section 2.2.3, an example of how multivalent might fit into this strategy is shown in Figure 15. As before, the first step is for a

bit level copy of the MPEG2 on the IMX tape to be made into a file. At first, there are no risks associated with MPEG2 encoding and this file is archived unchanged (1). However, imagine after a time that MPEG2 becomes obsolete and the format is considered at risk. Instead of either decoding or transcoding, the archive decides to port their existing MPEG2 decoder into Multivalent, or finds an alternative MPEG2 decoder that can be used in Multivalent (2). This allows them to keep the files in their original format and still access the contents. Eventually, when storage costs permits, the files are decompressed (3).

Figure 15 IMX video preservation example extended to include Multivalent
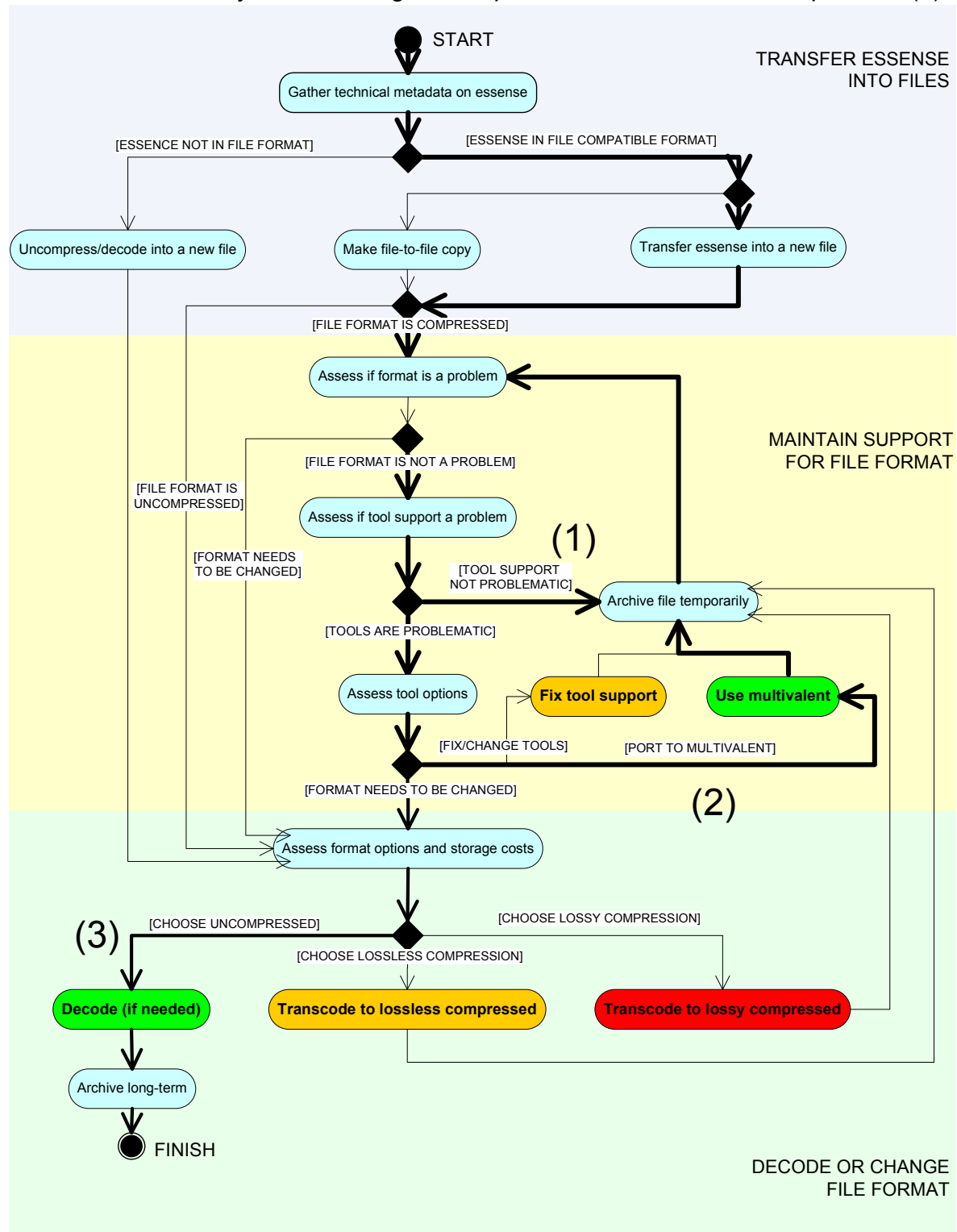
2.4.4.   *Multivalent maintenance lifecycle*

Whilst use of multivalent may provide an alternative to migration of video formats, it doesn't stop the need for migration altogether. Instead of migrating file formats for video, multivalent requires that the Virtual Machine (VM) and potentially the multivalent software (decoder) need to be migrated instead.

For example, if the platform on which multivalent is running (e.g. a windows or linux server) is upgraded then a new version of the VM may be needed to run multivalent. Alternatively, the VM could become obsolete, e.g. it is no longer supported by the vendor. Both cases would require a migration to a newer version of the VM.

For Multivalent as used in PrestoPrime, migrating virtual machines means new migrating to new versions of the Java Virtual Machine (JVM). Theoretically this very simple – you replace the old JVM with a new one and away you go. However, whilst that approach has generally worked for most java applications written to date, over the very long timescales of video preservation, various problems may occur:

Software applications written for one version of a JVM may not always work on a new JVM. Reasons include:

- Functionality in the old JVM is no longer supported in the new JVM, i.e. bits of the Java programming language are considered obsolete by the community. This is handled through the Java deprecation process. Obsolete language constructs get marked as deprecated which means that whilst they are still supported for now, they shouldn't be used as support will at some point be removed – although in the case of Java there is debate on when this should happen. In some cases, deprecated methods throw run time errors. There is a long list of deprecated Java language constructs (e.g. error handling, character sets, dates, AWT user interface etc.)

- The new JVM behaves slightly differently compared to the old JVM even though the functionality is nominally the same. In particular this can happen for JVMs implemented by different vendors (Sun are not the only people to supply JVMs, e.g. so too do IBM). If the multivalent software is coded on one JVM in a way that makes it dependent on the specific behaviour of that JVM, even if this is not intentional, then it might not function correctly on someone else's implementation of the JVM, or even on a different version of the JVM from the same vendor.

- Security is tighter in the new JVM, so things that worked in the old version of the JVM, e.g. access to resources esp. over a network or when using plug-ins or browsers, are denied.

These are reasons why a 'binary' version of multivalent might cease to function after the JVM on which it runs has been upgraded. These problems are not just specific to the multivalent software itself, but apply equally to any third-party code libraries used by multivalent.

If a problem does occur when migrating a JVM, then some of the options to solve the problem include: (a) updating multivalent; (b) alter the JVM so it behaves the way it did in previous versions (not a task for the faint hearted, but possible nonetheless as the JVM

code from Sun is available as source); or (c) try to find a way of keeping the old JVM running, e.g. through platform emulation.

None of these are attractive – especially updating the multivalent source code as this is the one thing that multivalent is meant to help avoid!

If there is a need to update the multivalent software, i.e. the decoders, then this means that the source code will need editing and then recompiling to create new binaries. This in turn adds further problems:

- The software development environment used to author and build the software may not work anymore, e.g. software development tools no longer work on current platforms.

- The software may not compile on the new JVM, e.g. due to deprecation warnings or more rigorous code checking by the compiler (e.g. non initialised variables).

- There may not be anyone familiar with the code which makes modification time consuming and error prone.

- Functionality previously provided by third-parties and accessed through libraries may now be part of the core language (XML parsing is a good example), which means any third-party libraries used are no longer available or are not supported.

If any third-party libraries used have licenses that mean they cannot be distributed as part of multivalent but instead need to be downloaded by the user directly, then this brings in further issues, e.g. whether to ignore these restrictions and archive 'everything' or to run the risk that they are not supported at some time in the future and can't be acquired when multivalent needs to be used.

Modern programming languages, e.g. Java, evolve rapidly. Changes to the language and JVM implementations happen on the timescale of a year or less - not the decades involved in AV preservation. This makes JVM migration a very real risk when using multivalent.

A diagram of the lifecycle for maintaining a multivalent approach is shown in Figure 16. Note that there is no end point to the diagram, i.e. the process is continual. Escape from the maintenance cycle is simply to stop using multivalent, for example by permanently decoding the files into an uncompressed form at which point the multivalent decoder is redundant.
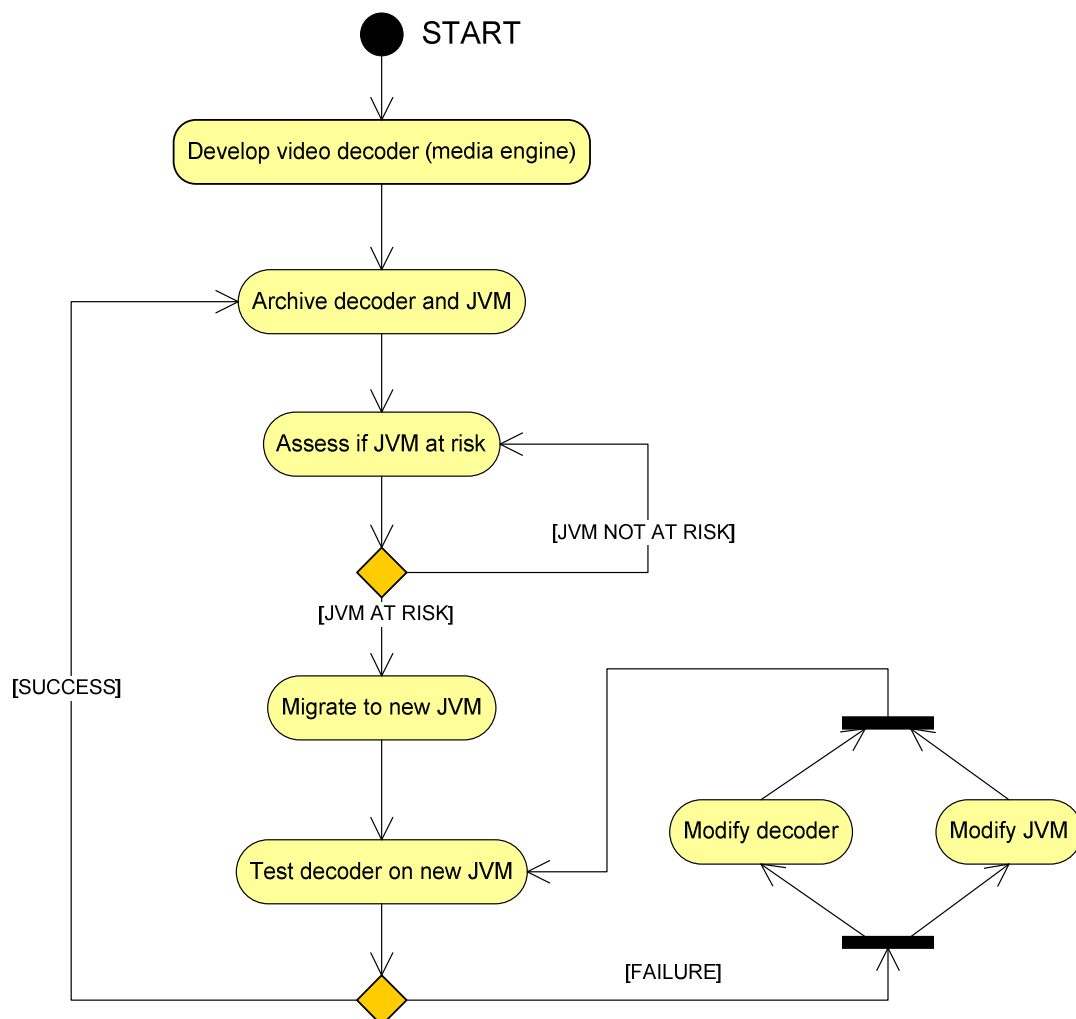
Figure 16 Lifecycle for maintaining multivalent decoders

2.4.5.   *Factors when considering multivalent*

Multivalent is attractive as it avoids the need for file format migration.

Migrating large volumes of video (10s or 100s of thousands of hours) between video encoding formats is time consuming, expensive and introduces risk of loss or degradation. Costs include human operators (e.g. for quality checking), software (e.g. transcoding), hardware (e.g. servers) as well as the overheads of planning, management, space, power etc... Avoiding these costs by leaving the content 'as is' and using Multivalent to access the content when needed is clearly very attractive.

Simply delaying migration can have major benefits too. For example, much of today's born digital content starts life as files in a lossy format (e.g. DV, IMX D10, XDCAM etc.). A typical preservation strategy is to ingest it in its original form and only change the format when the original format becomes obsolete, i.e. it is temporarily archived 'as is' until there is a risk of loss than means it needs to be migrated. At this point the content is ideally converted into a lossless format to avoid any further losses during transcoding. If budget permits, the format chosen is uncompressed (e.g. as defined within the SDI standards), not just lossless. This minimises the need for future transcoding and also makes the content more resilient to lossy storage systems (bit rot). Using Multivalent to defer a migration could make all the difference in terms of being able to afford uncompressed or

lossless compression. The rapid and ongoing rate at which storage/processing/networking costs fall mean it is likely that strategies currently used for audio , i.e. no compression, will be viable for storing today's video in 5 – 10 years time. Deferring migration saves cost (time and effort of people to do QA etc. as well as infrastructure costs), but the delay might not need to be for that long!

In most multivalent scenarios, there is a need to get the output of the multivalent system into another system and not just to be able to view content it in some form of browser. This 'other system' might be writing the content back to storage as new files in a new format, or it might be exporting the content out into an editing system so it can be repurposed/reused.

However, whilst multivalent has several plus points there are also other factors to consider.

Multivalent is in effect decoding on demand. If there are a large number of users of archive content, or the content is used frequently, then substantial infrastructure will be needed to do this decoding on demand. This is especially true in situations where high latency cannot be tolerated, e.g. content is needed in minutes rather than hours.

The 'decoding on demand' aspect of multivalent requires that performance as well as functionality will need careful consideration.

- In scenarios where multivalent is used to support browsing or viewing, then real time playback of high resolution video is likely to be required, e.g. HD, 2k or even 4k. This is not a trivial problem in Java. The file sizes and data rates can be huge. For example, file sizes of 100GB or more and bit rates of 200MBit/sec or higher are not uncommon for HD video (considerably more when uncompressed).

- In scenarios where multivalent is used to decode legacy formats 'on demand' then requirements for data volumes, throughput and latency all need to be considered. For example, imagine a collection of 100,000 hrs of video[67] in a legacy format encoded at 50MBit/sec[68]. If 25% of this content has to be decoded at least once a year[69] then that equates to a throughput of over ten concurrent real-time decodes running 24x7. The peak load could be considerably higher. For HD content with high bit rates, e.g. 440 or 880 MBit/sec, the problem is an order of magnitude worse. Here latency will be a major issue as users of the content will want access to it within minutes/hours and not days/weeks. In other words, access that is faster than real-time may be required.

- In scenarios where multivalent is used to delay format migration, but not completely avoid it, then there will come a point in the future where a whole collection will need to be migrated 'en masse', e.g. because it becomes viable to store it all in uncompressed form. This might mean using multivalent to decode say 100,000 hrs of video in the timeframe of weeks or months, i.e. massively parallel transfers or much faster than real-time decoding. Arguably when it is viable to do mass

---

[67] A large archive could have over a million hours of AV content, so 100,000 hrs in some cases is actually quite a small collection. For example, the BBC D3 transfer project is dealing with approx 100,000 hrs of D3 tapes (although not all will be migrated).
[68] e.g. SD video using DV in AVI files, or SD video with MPEG-2 encoding
[69] 25% of the BBC archive is accessed each year, with some items used more frequently.

decoding because storage is cheap enough, then the computing power of the computers available at that time will make very fast decoding possible, but careful analysis is required and there is the risk that an archive becomes locked into multivalent because mass decoding is not feasible in sensible timescales and decoding through multivalent is the only option (e.g. high performance hardware based decoders are no longer available).

2.5. Modelling the cost of preservation and access

There have been several attempts to produce cost models for digital preservation, including for audiovisual content as part of the PrestoSpace project, much of which is documented on the project's preservation website[70] and wiki[71] and is supported by a series of calculators (cost of storage and access)[72]. Several relevant and detailed presentations were made about the issues of cost[73] and building business cases[74] for audiovisual preservation and access at a TAPE workshop in 2008.

More general is the work of the LIFE project[75] coordinated by the British library. This has developed a detailed lifecycle for digital preservation and then developed methods to estimate the costs of each stage in the lifecycle over time.

The basic elements of the cost model are:

$$L_T = Aq + I_T + M_T + Ac_T + S_T + P_T$$

Figure 17 LIFE cost model (reproduced from How much does it cost? The LIFE Project - Costing Models for Digital Curation and Preservation[76])

Where:

L_T = Total cost, Aq = Acquisition cost, I = Ingest cost, M = metadata cost, Ac = Access cost, S = Storage cost, and P = preservation cost. The subscript T means that costs have to be calculated over the lifetime of the items being preserved.

The particularly valuable outputs from LIFE are the series of case studies[77] (web archiving, e-Journals, newspapers etc.) from the partners involved that include detailed spreadsheets implementing the LIFE model and provide real-world worked examples of what the costs of preservation really are. The LIFE approach is founded on considering cost over time, e.g. for 5,10 or 20 year periods, with the result that the LIFE model and examples include activities such as migration and time varying costs such as storage, both of which are very relevant in the audiovisual world.

---

[70] http://digitalpreservation.ssl.co.uk/
[71] http://wiki.prestospace.org/pmwiki.php?n=Main.PrestoPRIME
[72] http://digitalpreservation.ssl.co.uk/about/56/59.html
[73] http://www.hatii.arts.gla.ac.uk/confweb/tape/matthew_cost_models.ppt
[74] http://www.hatii.arts.gla.ac.uk/confweb/tape/tape-outputs.html
[75] http://www.life.ac.uk/
[76] http://liber.library.uu.nl/publish/issues/2007-3_4/index.html?000210
[77] http://www.life.ac.uk/2/documentation.shtml

In the context of audiovisual preservation, the elements of the LIFE model are directly applicable. The LIFE approach provides a helpful and structured way to do cost calculations for digital audiovisual preservation and access projects. The nature or audiovisual material is such that specific parts of the LIFE cost model can dominate others, for example the storage costs of the very large data volumes that can be involved, the access costs for online streaming of video content, or the quality control costs of manual content appraisal for example after digitisation or transcoding. However, all the elements of the LIFE model do apply in the audiovisual domain. The specific cost areas likely to be most significant for AV material are explored in more detail below.

Acquisition	Ingest	Metadata	Access	Storage	Preservation
Selection (Aq1)	Quality Assurance (I1)	Characteri-sation M1)	Reference Linking (Ac1)	Bit-stream Storage Costs (S1)	Technology Watch (P1)
IPR (Aq2)	Deposit (I2)	Descriptive (M2)	User Support (Ac2)		Preservation Tool Cost (P2)
Licensing (Aq3)	Holdings Update (I3)	Administrative (M3)	Access Mechanism (Ac3)		Preservation Metadata (P3)
Ordering & Invoicing (Aq4)					Preservation Action (P4)
Obtaining (Aq5)					Quality Assurance (P5)
Check-in (Aq6)					

Figure 18 Breakdown of cost elements in the Life model (reproduced from How much does it cost? The LIFE Project - Costing Models for Digital Curation and Preservation)

Before looking at specific cost areas for audiovisual preservation and access, it is worth noting that the LIFE model is just one of several attempts to look at the overall lifecycle of digital content. Looking at the lifecycle and then considering the cost as a function of time for each stage provides a methodical approach to calculating total costs.

Other examples of lifecycle models include the DCC (Digital Curation Centre) Lifecycle model as shown in Figure 19.
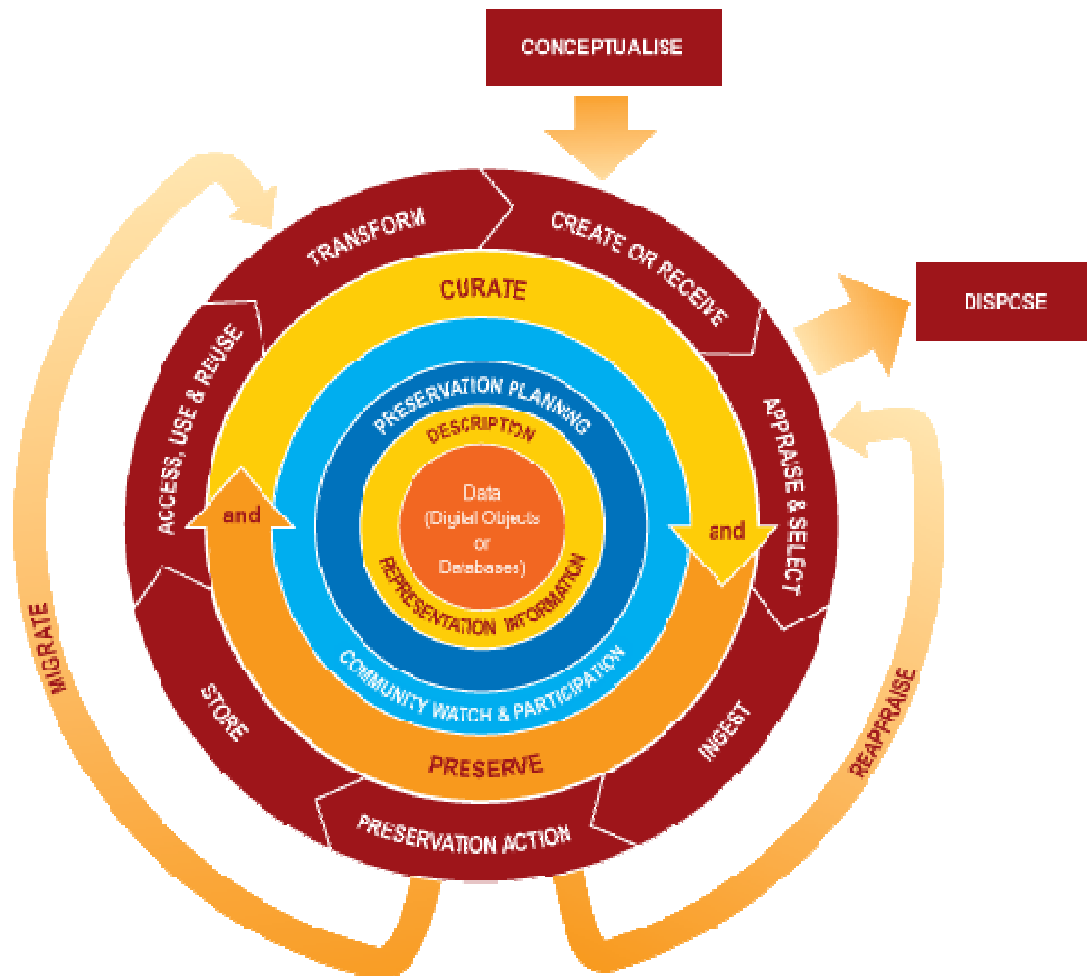
Figure 19 DCC Lifecycle model (reproduced from the DCC website **http://www.dcc.ac.uk/lifecycle-model/**)

The DCC model contains a similar set of stages in the lifecycle to the elements in LIFE and highlights the need to consider planning, community watch and metadata aspects at all stages in addition to the digital objects themselves.

Whilst the DCC and LIFE models above are relatively generic, similar lifecycle models exist that describe audiovisual media, for example as shown in Figure 20 for broadcast media as it goes through the lifecycle of production, post production, delivery and consumption. This lifecycle model emphasises the central role the archive plays in all aspects of the media lifecycle and hence the need to consider the archive in this wider context when calculating costs. For example, knowing the workloads placed upon the archive infrastructure in order to serve the needs of production is an important part of capacity planning for the archive infrastructure (people, servers etc.) which in turn has a major bearing on costs. This highlights the need to consider the community that creates and consumes content in the archive when considering costs.
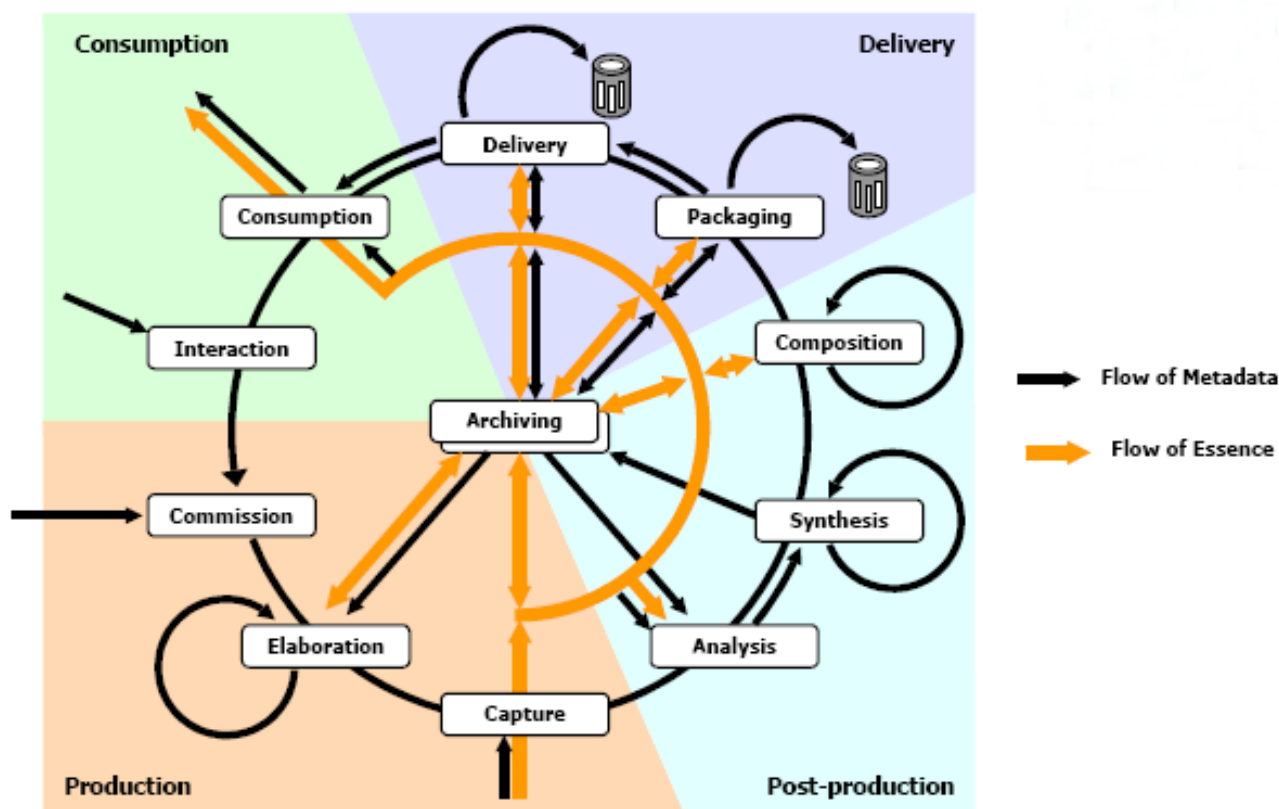
Figure 20 Broadcast media lifecycle (reproduced from a presentation by Tom Coughlin on his Report in 2009 on Data Storage for professional Media and Entertainment[78])

The Interim report of the Blue Ribbon Task Force on Sustaining the Digital Investment[79] provides a further set of examples on the cost of preservation (economics), including real numbers for categories such as staff and equipment. Again the value here is the input it gives into what cost categories need to be considered in the total cost of preservation and access.

Likewise, The Preserving Public Television Project[80] on Strategies for Sustainable Preservation of Born Digital Public Television[81] also provides considerable detail in particular through a case study on the costs of a prototype digital repository built by NYU Digital Library Technology Services. This case study is notable as it looks in detail, with numbers, on the staffing requirements. Experience suggests that labour costs can be significant in preservation projects, yet few archives have undertaken large scale and modern digital repository projects, and hence information in this area is often hard to find. This makes the detailed figures from NYU particularly valuable.

Finally, as an example of cost modelling specifically in the audiovisual preservation area, "The Digital Black Hole" by Jonas Palm from the Riksarkivet[82] provides a quite well known

---

[78]

http://www.tomcoughlin.com/Techpapers/Data%20Storage%20for%20Professional%20Media%20and%20Entertainment,%20060809.pdf
[79] http://*brtf.sdsc.edu/biblio/BRTF_Interim_Report.pdf*
[80] http://www.thirteen.org/ptvdigitalarchive/
[81] http://www.digitalpreservation.gov/
[82] http://www.tape-online.net/docs/Palm_Black_Hole.pdf

example of how to cost long-term digital storage in terms of staff, equipment, facilities etc…

2.5.1.   Cost of Analogue Film

The 'keep film as film' approach has been analysed by the motion picture industry itself, in their report The Digital Dilemma[83]. Their overall conclusion was that storing film as film was 11 times cheaper than storing it digitally.

Specifically: "Economic models comparing long-term storage costs of film versus digital materials show that the annual cost of preserving film archival master material is $1,059 per title, and the annual cost of preserving a 4K digital master is $12,514"

The $1,059 comes from: "*a monthly cost of 40 cents per 1,000 foot film reel in preservation conditions plus the amortized cost of film archive element manufacture*". Three *archive elements* are produced: a YCM, a finished negative, and an interpositive – at a cost of $80,000, which is 76% of the total 'century cost' of $105,900 – and is of course incurred at the beginning.

Further: "access costs are not included" in the Academy study.

A different study on precisely the same issue – keeping a "cinematic work" for 100 years, was performed by Sun Microsystems[84] and got a different result – primarily because that study did include access costs, and had much lower estimates for digital storage costs. Sun estimate the cost for ONE access as follows: "If we want to repurpose this movie only one time in the century, we must re-scan the colour separations ($65,000). If any restoration work is required, that is an additional cost."

The Sun model also saved all the original camera negatives, which added $70k to the total bill. Taking that out, the Sun estimates just for film-based *storage* (not access) for a century was $60k. This is 40% less than the Academy figure, but it is useful that these two 'competing' estimates are within a factor of two regarding analogue costs.

The point at issue is: what is the use model? If, following the Academy, the model is low use for a century, then the cost is in the region of $60k to $100k for film-based storage – and the bulk of that cost is paid up-front in production of film elements – plus a per-access cost of about another $65k.

The BBC runs a broadcast archive, where content tends to be reviewed on a five-year cycle, and where usage is the entire justification for the archive. As with the rest of broadcasting, these archives contribute heavily to everyday output: 2000 items per day come out of the BBC's audio and video archive, and about 20% of the content is accessed per year. That means each item is accessed, on average, once in five years, or 20 times in a century.

---

[83] http://old.oscars.org/council/digital_dilemma/index.html
[84] *Archiving Movies in a Digital World*, Dave Cavena et al, January 10, 2007;
http://sun.com/storagetek/disk_systems/enterprise/5800/ArchivingMoviesinaDigitalWorld.pdf

For the BBC usage, the analogue storage approach would cost roughly 21 times $65k = $1,365k. That compares with the Academy's estimate of digital costs as being 100*$12.5k = $1,250k.

Clearly access makes an enormous difference to the 'save as analogue' model, because one access (one re-scan of the colour separations) costs about as much as storage over the entire century. The choice between the analogue and digital options thus depends entirely on the use model: even one use (specifically, the re-scanning attendant to that use) costs as much as a century of shelf-based storage.

2.5.2.   *Cost of Digital Storage*

The conventional IT approach to storing data files, including audiovisual material, is to use solutions based on Hard Disk Drives (HDD) or data tape, including a combination of the two, for example in a Hierarchical Storage Management (HSM) system[85]. Deliverable D2.3.1 of the PrestoPrime project describes in detail the storage system architectures used by audiovisual archives, with specific examples from the project partners.

The question is what does this type of storage cost and how does that cost change over time?

There have been several studies that report the cost of digital storage, including the Total Cost of Ownership (TCO). These include the San Diego Super Computing Centre (SDSC)[86], an extensive Google white paper on warehouse scale computing[87], the breakdown in the Digital Black Hole report described earlier, and from those who have tried to push the boundaries on low cost mass storage through the use of commodity components, e.g. BackBlaze[88]. All look at the cost of network accessible managed storage, e.g. HDD servers or tape robots. There are also commercial storage services that publish their rates, for example Amazon S3.

Before getting into the details, the headlines results are clear:

1PB of data online stored using HDD has an annual TCO of $1M

Data tape (in robots) is approximately a third of this cost.

1 hour of HD uncompressed is approximately 1TB (1080p50 at 1.5Gbit/sec). The annual storage cost (for the first year) will be $1000. 1 hour of SD uncompressed is approximately 100GB (200MBit/sec) and hence costs $100 for a year. A large archive that has 100,000 or even 1,000,000 hours of content is therefore immediately faced with an annual storage cost in the tens of millions of dollars.

---

[85] http://en.wikipedia.org/wiki/Hierarchical_storage_management
[86] Moore, R. L.; D'Aoust, J.; McDonald, R. H.; and Minor, D. (2007). Disk and Tape Storage Cost Models. In *Archiving 2007. http://chronopolis.sdsc.edu/assets/docs/dt_cost.pdf*

[87] Barroso, L. A. and Holze, U. (2009). *The Datacenter as a Computer: An introduction to the desin of warehouse-scale machines.* Google Inc. Synthesis Lectures on Computer Architecture no. 6. published by Morgan and Claypool.
http://www.morganclaypool.com/doi/pdf/10.2200/S00193ED1V01Y200905CAC006?cookieSet=1
[88] http://blog.backblaze.com/2009/09/01/petabytes-on-a-budget-how-to-build-cheap-cloud-storage/

The exact costs vary depending on how the calculation is done and where the supporting cost data comes from.

The Academy estimate $12.5k per film per year are based "on an annual cost of $500 per terabyte of fully managed storage of 3 copies of an 8.3 terabyte 4K digital master". But Sun estimate the cost, based on economies of scale in an overall archive with at least 2000 titles, as $73k per title per century. The Sun estimate is 6% the size of the Academy estimate! One difference is that Sun estimate costs for an 'object' that takes up 100TB, while the Academy was 'only' 25 TB (three copies of a 8.3TB master).

How are engineers and archivists suppose to resolve a situation in which major authorities in the field disagree by a factor of 16 in estimating digital storage costs (by a factor of 64 if the different object sizes are included)?

We aim to unpick some of the details and provide some clarity. There are two key issues.

1) The total cost of managed storage. This exceeds the cost of raw media by a factor of 10.

2) The change in storage costs over time. Storage costs halve every 3 years or faster.

The SDSC study provides a cost breakdown for both data tape and HDD storage, which include media cost, capital costs, space, facilities, maintenance etc. This is shown in Table 3. This shows that hard disk media cost is only approx $1/3^{rd}$ of the total cost, with tape media cost being only 20% of the total cost of storage.

	SATA Disk Storage (1.8 PB)		Archival Tape Storage (5 PB)	
	$/TB/ yr	% of total	$/TB/ yr	% of total
Disk/cartridge media cost (annualized)	535	36%	100	20%
Other capital costs (annualized)	235	15%	165	33%
Maintenance & license costs	230	15%	110	22%
Facilities Costs - space, utilities	160	11%	25	5%
Maint./System Admin Labor (@ $150K/FTE)	340	23%	100	20%
Total Cost	1500		500	

Table 3 Cost breakdown for data tape and hard disk storage (reproduced from SDSC 2007 paper on Disk and Tape Storage Cost Models).

A similar analysis can be done based on storage services, e.g. Amazon S3. Storage costs for large volumes of data are approximately $0.10 cents per TB per month, i.e. just over

$1000 per PB per year. Amazon use hard disk storage. The raw media cost of a 1TB SATA drive is less than $100. This gives a 10:1 multiplier for total cost. It also not strictly speaking correct to compare today's media cost with the Amazon charging rate since they will of course be amortizing disks over several years and hence part of the Amazon cost is covering servers and media purchased at a time when they were probably twice the price. Amazon also has at least two disk copies of the data. Therefore, the effective multiplier is 5:1 or less, which is consistent with the SDSC figures.

The Google warehouse scale computing study doesn't differentiate between media cost and capital costs of the servers or robots that the media is held in (mostly because the Google study focuses on computing resources rather than storage resources). It does however look in detail at the power and cooling costs, which is a substantial fraction and a key area to address in driving long term costs down (the cost of electricity will surely rise significantly over time in contrast to rapidly falling hardware costs). The Google study considers two case studies with different types of storage server in terms of cost vs. power consumption. The server component (which includes media) is between 25% and 50% of the annual TCO. This is consistent with the SDSC study.

For disk based storage the cost can vary enormously depending on the speed of access required to the data on the disk (for example, consider high performance storage clusters using enterprise disk vs. low cost SATA disks in a MAID array). This difference is illustrated nicely in the BackBlaze analysis shown in Figure 21.
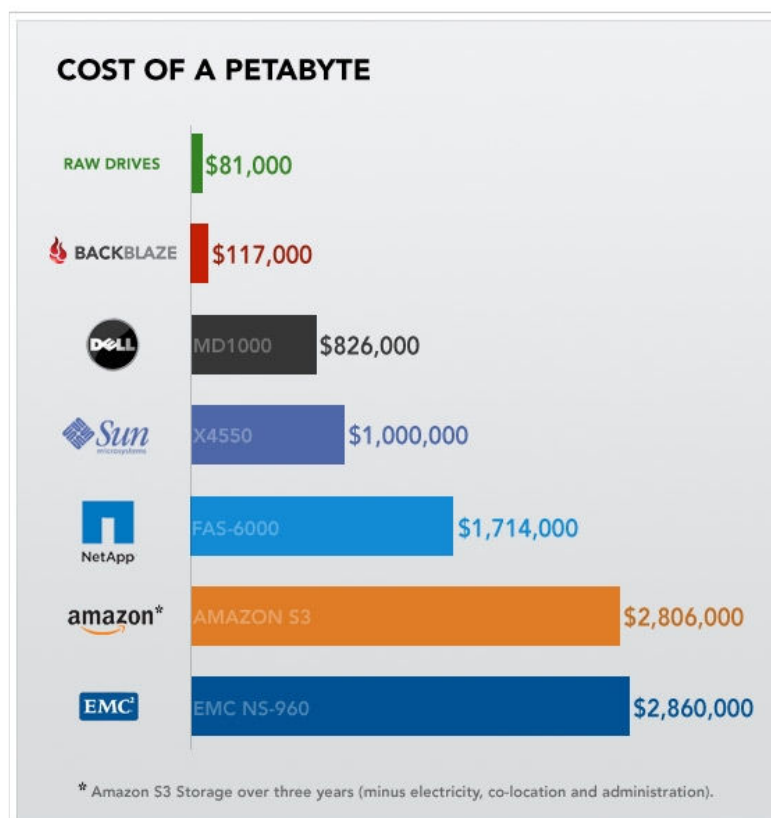

Figure 21 Storage cost comparison for different HDD solutions (reproduced from BackBlaze analysis[89])

---

[89] http://blog.backblaze.com/2009/09/01/petabytes-on-a-budget-how-to-build-cheap-cloud-storage/

BackBlaze are pushing the boundaries of how the overhead of a storage server can be made as low as possible. They do this through design and build of their own storage servers instead of purchasing off the self units from established suppliers or by using storage as a service from third-parties e.g. Amazon. The viability of this DIY approach for long term archiving is unproven, and there are of course a new set of costs, e.g. ongoing design and development of new servers, but BackBlaze does serve to illustrate what is theoretically possible when driving down costs for HDD storage.

The BackBlaze study factors out facilities cost, e.g. overhead of having storage in a data centre, but these are easy to factor back in by using the breakdowns from Google, SDSC and others. The BackBlaze study does however show that the total cost of storage can exceed 10 times the raw media cost, at least when considering today's media cost.

When looking at the long term costs of storage it is essential to consider the rapid rate of technical obsolescence and the limited lifetime of both media and servers. More detail on this can be found in PrestoPrime deliverable D3.2.1 which looks at the threats to audiovisual content from the use of mass storage technology. Roughly speaking, HDD have a life of 3-5 years and data tape 5-7 years. For HDD it is the media lifetime (the spinning disks) that is the limiting factor. For data tape, e.g. LT04, the media (the tape) is much more stable, but new players make old tapes obsolete and unreadable after just a few generations (the LTO roadmap says tape drives for generation X should be able to read/write tapes of generation X-1, and read tapes from generation X-2, but doesn't require the drive to read tapes of generation X-3. There is a new generation every 2-3 years). Therefore, storage media and servers are typically amortised over a 3-5 year period.

At the same time the unit cost of storage (HDD or tape) is falling rapidly. This follows a trend similar to Moore's law as shown in Figure 22. A similar trend exists for data tape, although it tends to be more a series of steps, e.g. following the LTO roadmap.
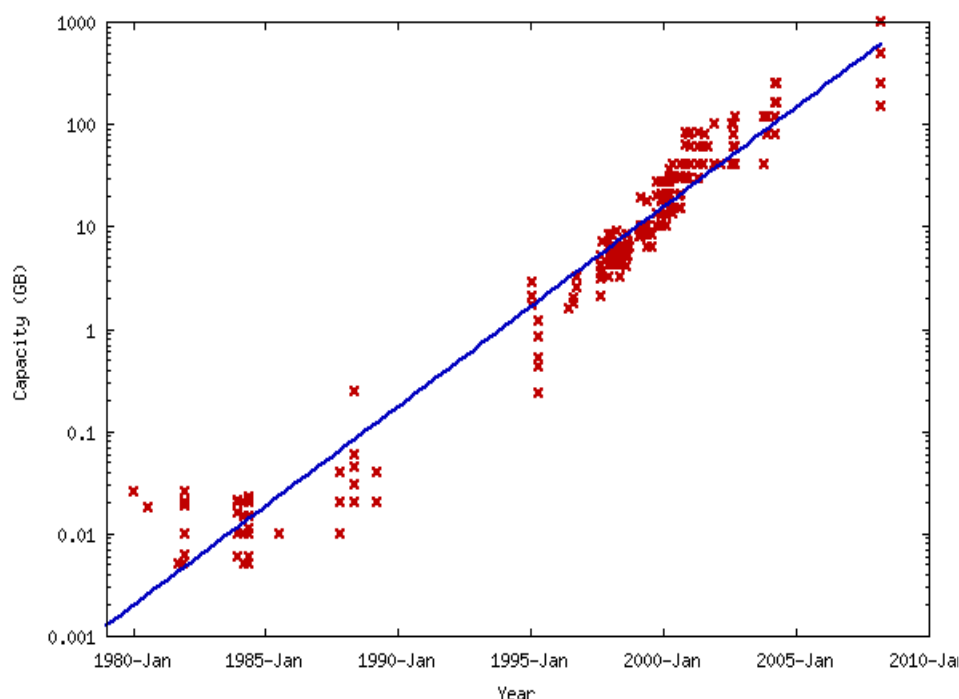

Figure 22 HDD capacity as a function of time

Roughly speaking the media cost for storing a TB of data halves every 18 months, which is a factor of 100 in a decade. Within three decades, if the trend continues (as it has for the last three decades), then cost will have fallen by a factor of 1 million. This means that the entire digitised content of today's audiovisual archives will fit on a single hard drive server[90]

The TCO of storage when provided by SDSC, Amazon or others doesn't fall as fast since the costs are amortised. These facilities, having been established for some time, and serving the needs of a community with ever growing volumes of data to store, necessarily have a collection of storage servers of varying ages. In this scenario, if a storage system is written off over three years and the cost of such a system halves every 18 months (approx 20% cost reduction per year), then the average cost of storage across an active collection of such servers at of varying ages will be double the cost of new storage.

The rate at which the TCO of storage falls doesn't appear to be quite as high as the rate at which raw media costs are falling. SDSC report that the cost of tape halves every three years with the rate for disk being somewhat faster. This is reflected in the rates they charge to customers ($1500 per TB per year on disk in early 2007, $1000 in 2008 and currently $650 as of the end of 2009[91]). Similar trends are seen in the Amazon S3 rate, which in 2007 when SDSC produced their analysis stood at approx $1800 per TB per year (see SDSC paper) and now stands at $1260 per TB per year (if you store over 500TB)[92] and considerably less if you have multiple petabytes to store with the extreme being only $660 per TB for over 5PB of storage.

It would appear that the real-world TCO for storage halves every 2 to 3 years, with disk storage falling faster than tape storage. Tape storage using the most recent SDSC figures is still half the cost of disk storage (and that includes 2 copies on tape against one on disk).

Knowing today's annual TCO for storage and the rate at which it falls then allows us to calculate the TCO for storage in perpetuity. The calculation is in Annex I of this report. The results are below in Table 4.

Number of years for annual cost to halve	Total cost for storage in perpetuity as a multiplier of the year 1 cost
1	2
2	3.3
3	4.9
5	7.7
10	14.9

Table 4 TCO for storage in perpetuity as a multiplier of today's annual storage cost

If the time for the cost of storage to halve is between 2 and 3 years, then the total cost of storage forever will be approximately 4 times today's annual cost.

---

[90] There is an estimated 100M hours of AV content in the worlds archives (UNESCO). If all of that was video at SD resolution (200MBit/sec) then uncompressed storage would require 10 Exabytes (10,000 Petabytes). This is 10 million times the size of the average HDD of today (1TB), so would all fit on 10 drives of the future in 30 years time, i.e. a most hard disk server.
[91] http://www.sdsc.edu/services/StorageBackup.html
[92] http://aws.amazon.com/s3/

Interestingly, if the cost of tape is falling slower than the cost of disk, which it is, and if this continues, which appears likely, then the long term cost of a disk based solution will be comparable to the long-term cost of tape. Of course, the year 1 cost will be higher and convincing the bank manager that this is the best option could still be difficult!

Going back to Sun's study, unfortunately it doesn't give a detailed breakdown of costs, but clearly $73k per century for a 100TB appears to be on the low side. Using the $500/TB/yr for tape from SDSC paper which gave 2007 costs at the time that Sun did their calculation, and then using a multiplier of x4 to give the in perpetuity costs, then 100TB for a century would be $200k not $73k

So, in summary, our simple set of rules for calculating long-term storage costs are:

1. Take today's cost for commodity storage media (SATA hard disk drives or LTO4 tapes) and multiply by 10 to get the total annual cost of storage using that media.

2. Repeat step one if you want multiple copies, e.g. disk and tape.

3. Multiply the result of step two by 4 to give the TCO for a lifetime of storage.

These rules only work for relatively large volumes of storage, e.g. 10s or 100s of terabytes and assume that the objective is long-term preservation, e.g. a decade or more.

For example, 1 hour of SD uncompressed requires 100GB of storage. A collection of 10,000 hours would require 1000TB. Unit media cost is currently approx $100 for a 1TB SATA drive and $50 for an 800GB LTO4 tape. One copy on disk and one copy on tape would have an estimated year 1 TCO of $750,000 and a lifetime cost of $3M.

At this point it is worth looking at some of the wider trends in storage.

Firstly, online storage requires power and cooling, especially for spinning disk. Power consumption for the average SATA drive is around 10W[93], higher at start-up, but lower when the drive is idle. A thousand 1TB drives therefore consume 10kW power. Further power is required by the servers in which they sit, e.g. to keep the fans running that cool the disks as well as for onboard electronics etc… The cost breakdown for the BackBlaze storage server suggests that up to twice the power is needed for the server as a whole as for the individual drives. That power requires a electricity supply at one end and will generate a heat load at the other, which in turn requires further infrastructure for cooling, e.g. in a data centre.

The ratio of total power required as input to a data centre to the total electrical power consumed by the equipment in the data centre (servers etc.) is called the Power Usage Efficiency (PUE) rating[94]. A typical value for a data centre[95] is around 2, with values as low as 1.2 possible[96] when taking extreme measures (e.g. locating the data centre next to geothermal or hydroelectric sources of power and where the climate permits free air cooling, e.g. in a desert at night or in the arctic). The converse can also be true for facilities

---

[93] http://ixbtlabs.com/articles2/storage/hddpower.html

[94] http://en.wikipedia.org/wiki/Power_usage_effectiveness

[95] http://www.google.com/corporate/green/datacenters/measuring.html

[96] http://www.datacenterknowledge.com/archives/2008/10/01/google-the-worlds-most-efficient-data-centers/

that either don't make use of well designed modern data centres or don't adopt state-of-the-art practices in house. Here a PUE of 3 would be more likely.

Therefore, 1PB of online disk has an annual energy consumption of between 100,000 and 300,000 kWh. Using an approximate energy cost of $0.10 per kWh, the annual energy cost will be $10k - $30k. This is significant, but not actually that large compared with the TCO for the storage, which is 20 times higher. With large modern data centres designed for megawatt power loads, the 10s of kW of power needed for storage servers is not a large burden either. All this suggests that, for now, power and cooling are not a major factor when considering storage solutions or locations.

However, other related activities in the production, post-production and distribution process can require significant compute resources (e.g. the 60 node server farm behind iPlayer). Major cost savings can be achieved by locating them in energy efficient data centres. In turn, due to the integral role that a modern archive plays in these processes, it may well be the case that the archive follows suit and also gets relocated.

The next trend worth looking at is the move towards solid state based storage, e.g. flash memory.

Solid State Drives (SSD)[97] are falling rapidly in price. Whilst touted mostly for their performance compared to HDD, from an archiving perspective SSD could be attractive due to lack of mechanical parts and potentially long data retention periods without the need to be powered on. However, the cost is still high compared with HDD, with the unit media cost being 10 times higher at present (approx $0.10 per GB for HDD and $1 per GB for flash, either in memory cards or SSDs). The need for less replication, less frequent migration, less power and cooling, and lower maintenance means that the TCO of flash based storage is more likely to be 5x that of HDD rather than 10x, but this is still prohibitively high for archiving. However, given the rate at which flash based storage is falling in price, it is certainly worth keeping a long-term eye on this technology.

The cost effectiveness of one type of storage solution over another, e.g. HDD vs. data tape, depends on the amount of content to be stored. For example, at one extreme of a very small archive, e.g. a few GB of content, 2 or more solid state devices, e.g. USB sticks would provide an adequate solution. Neither tape nor disk would be as cost effective. At the next level up, hard drives become viable but data tape is still relatively expensive due to the need to purchase a tape drive in order to read/write the media. At the other extreme of very large archives, a data tape solution can easily be more cost effective than hard drives.

- The tipping points depend on the size of the archive and moves from solid state→ HDD → data tape as the archive gets bigger, with mixtures of technology viable in-between, e.g. frequently accessed content on HDD with the rest on near-line data tape.

- The tipping points will move over time, i.e. what makes sense to store on data tape today could become viable to store on hard disk in say 5 years time.

---

[97] http://en.wikipedia.org/wiki/Solid-state_drive

2.5.3.   *Cost of Tapes on Shelves*

If, as shown in the previous sections, that the lifetime TCO for online storage using tape or disk is approximately 40 times the raw media costs, then an obvious question is why not adopt a 'media on shelves model'.

Taking this approach for HDD is a recipe for disaster given the real-world annual failure rates of these devices (see PrestoPrime D3.2.1 for a review of this area) where 2-4% of drives failing each year is not uncommon. Compounding the rate over say 5 years gives a high probability that a failure will occur. This is of course why RAID and other techniques have been developed. Very long-term un-powered storage of drives is likely to cause problems due to mechanical issues. However, with periodic spinning up of drives it can be possible to reduce the rate of failure compared to continually operating drives by a factor of 2 or more[98,99] which leads to the use of MAID[100] (Massive Array of Idle Disk) approaches, e.g. as implemented by COPAN[101]. This is however a storage system and not a 'disks on shelves' model.

- Here our recommendation is simple: don't put un-powered HDD on shelves, inc. USB attached external hard drives.

'Tapes on shelves' is a common approach for archiving, in particular for both analogue and digital audio and video tape. When transitioning to files, an obvious choice would appear to be a data tape on shelves model.

Whilst data tape itself should be reliable enough as a storage media, the need to migrate every 5-7 years due to obsolescence of the drives requires all the tapes to be taken off the shelf and the contents transferred to a new tape. This necessarily involves a manual step. Likewise, if there is the need to access the content on the tape, then manual retrieval is necessary. Neither costs apply to automated storage using tape robots.

Take the BBC as a hypothetical example. The BBC D3 project is transferring 100,000 hours of D3 digital video tape content into files on LTO data tape using uncompressed SD as a file format (100GB per hour of programme material). The storage requirement is therefore approximately 10,000 LTO4 tapes.

20% of the BBC archive is accessed every year, so over a 5 year period we can expect every item to be accessed on average once (some more and others not at all). Moving the items from 100,000 D3 tapes (1 hour per tape) to 10,000 LTO4 tapes means there is approx 10 items per data tape. Therefore, even though each item is accessed on average every 5 years, this requires all the data tapes to be retrieved each year. Each tape also needs to be accessed once every 5 years for migration. So, the average number of accesses per tape is 6 times in 5 years.

The access requirement, including migration, is therefore 120,000 tape retrievals over 5 years, or approximately 70 per day. This would appear to be a full time job for at least

---

[98] http://dsp.ucsd.edu/~jfmurray/publications/Hughes2004.pdf
[99] Discussions between the author and COPAN staff (manufacturers of MAID technology) suggest that periodic spinning up of disks that are otherwise idle can extend the life up to 3x beyond the manufacturers figures. Given that field studies of continuously spinning disks show failure rates much higher than manufacturer claims, then MAID using disk aerobics can provide up to 10x extension in life.
[100] http://en.wikipedia.org/wiki/MAID
[101] http://www.copansystems.com/

three people given the need to retrieve and load each tape into a drive and transfer the content each time an access happens.

Assuming LTO4 tapes at £30 per tape, the total cost of the tapes is approx £300,000. Three people working full time might cost £150,000 when office space etc. is included. Therefore, the people costs of the tapes on shelves model will equal the media cost in only 2 years. The people cost will be the dominant factor over a longer period, e.g. 10 years.

The cost of data tapes in a robot has an estimated lifetime cost of $12M using current SDSC rates, i.e. approx £8M. This includes two copies. The 10 year cost for two copies using a tapes on shelves model as described above is approx £2M (ignoring inflation in staff costs and the extra space costs of shelves compared to a robot). Tapes on shelves is ¼ of the cost of tapes in a robot. If each video item was accessed once a year rather than once every 5 years then the cost of the two approaches would be the same.

This simple example hopefully shows that a 'tapes on shelves' model, whilst superficially very attractive due to low equipment cost, has significant long-term staff costs that need careful consideration.

2.6. Cost of Access

2.6.1. Analogue Film

Comparing the Academy costs for preservation of analogue film for a century with the costs of storing the same film digitally results in similar numbers:

- century cost of analogue film, per title: $105k
- century cost of preserving content on digital storage, per title: $125k (=$12,514x10, rounded)

Superficially, these seem comparable – until the cost of regular access is included. If scanning is indeed about $60 per title as per the Academy study, then just one access tips the balance. The BBC's 20 accesses per century would make the digital approach far cheaper (making analogue film 10 times more expensive than digital, rather than vice versa).

2.6.2. Online Access

Online access to AV content whether for internal, commercial or public access reasons is a major driver for preservation (see case studies in Section 1). The cost of access needs to be built into the overall cost model and can be very significant. Underestimating the costs can easily cause fragile business models surrounding online access to fail[102].

One simple way to compare the cost of access to the cost of storage (which is part of the cost of preservation) is to use current cost data from cloud service providers.

The Amazon S3 storage service[103] charges storage on a per TB per month basis and likewise for access to that data (both the initial transfer to Amazon as well as accessing data already there). The rates do fall as the data volumes being stored or accessed go up,

---

[102] http://www.businessinsider.com/2007/9/analyzing-econo
[103] http://aws.amazon.com/s3/#pricing

but roughly speaking the rates are currently (December 2009) $0.10 per TB per month for storage and $0.10 per TB per month transferred to or from that storage.

It is immediately obvious that if the data is accessed once a month then the network costs are comparable to the storage costs. If data is accessed much more frequently, then network costs can outweigh storage costs by a considerable margin. So, for example, for the INA 'archives for all' campaign (see Section 1.1.3) in November 2008 they were seeing 1.5 million videos streamed per month from 100,000 items of content. Each item is therefore accessed, on average, nearly 200 times a year. The access cost using the Amazon charging model is 15 times the storage cost.

This 10:1 ratio between network and storage costs for online content is explored in some detail in "Economics of Online Video 2: Unit Cost Structure"[104] including a set of spreadsheets that also look at the return from content sales and whether this can cover the costs. In the baseline analysis[105], there is a 8:1 ratio of bandwidth to storage costs for online video with 1000 accesses per item per year.

It should be noted that 'heavy access' scenarios are typically for 'browse resolution' content, e.g. web quality and not for master quality material. Here the storage costs of the digital master can outweigh the access costs. For example, only 20% of the BBC archive is accessed each year (which is still a lot considering the BBC has several million items), but it does mean that the storage component of the cost will dominate the access component – which is why the BBC and archives like them still tend to use a 'tapes on shelves' or 'tapes in a robot' approach as these have the lowest storage costs.

However, with increasing reuse of high quality archive content the trend will inevitably be towards HDD based storage which facilitates rapid online access, and away form the relatively longer latencies of data tape based storage.

2.6.3. *Case Study: the cost of access for the BBC iPlayer*

It is difficult to get exact BBC costs on storage and distribution, because costs of specific services come from commercial contracts, so that data is sensitive. However two general points can be made:

1) Theory: to make 'an item' available for electronic access, only one copy of that item has to be stored by the access provider, while each viewer requires, effectively, his or her own individual copy to be distributed. Therefore anything that is reasonably popular will generate a lot of traffic, but no additional storage. Therefore the *marginal* cost – the extra cost per viewer of adding 'views' to an existing service – is zero for storage[106], but costs the "network cost" for an item of that size, per new user, which can be estimated based on existing cost models, e.g. Amazon S3.

---

[104] http://www.businessinsider.com/2007/9/economics-of-on
[105] http://spreadsheets.google.com/pub?key=pYFPEp3S18PRSx2k2tnDrJA
[106] This isn't strictly true because when number of accesses goes up so does the required performance of the storage system that services this access. For example, a tape library wouldn't be a suitable choice for a large number of online users, where a HDD solution would be more appropriate, but for infrequent access, e.g. once a week, then a tape library could be more than enough.

2)   Practice: the BBC annual report[107] shows that TV distribution in 2008 cost £104.8 million, and online (web) distribution cost £15.4 million. What did the storage cost? The figures are not in the annual report, but the commercial rate for such storage can be estimated as "under £10 thousand"[108]. So total web distribution cost was about 1500 times higher than (estimated) storage cost.

Interestingly, an analysis of the cost per viewer of online streaming vs. traditional broadcast using radio transmitters concludes that online access can be 16,000 times[109] as expensive!

Further details of the huge data volumes delivered through iPlayer are available in an article by Nate Lanxton[110], with current (May 2009) estimates being 7PB streamed per month, which would equate to approx £5M per year at Amazon S3 rates for network traffic and is consistent in order of magnitude with the BBC annual report.

2.7.   *Key Findings and Conclusions*

The issues of preserving audiovisual content in a given file-format are complex, although the overall long-term strategy is simple: uncompress.

Selecting a suitable preservation strategy requires consideration of the issues of sustainability, in particular tool support, the ways in which digital content can be stored and the Total Cost of Ownership over time of this storage, and how the content needs to be accessed and used.

Our simple recommendations here are to adopt an open wrapper format (e.g. MXF), use uncompressed audio bitstreams (e.g. BWF or AES3) and then do the following with the video:

- If the video content is already in digital file format and is sustainable then you should keep it unchanged.

- If you have to change format and can afford uncompressed storage, then this is the first choice.

- If you have to change format and can't afford uncompressed, then lossless compression using an open format, e.g. JPEG2000 is the next best thing.

- If you have to change format and can't afford lossless compression, then choose a lossy format that is long-lived and open with a quality level that matches the source material.

---

[107] http://www.bbc.co.uk/annualreport/
[108] Estimating the cost of BBC iPlayer storage: Assuming iPlayer holds about 1000 hours of content, at a coding rate of 500kb/s, the storage would be 250 GB (because it would be approximately 250 MB/hr). At Amazon prices of $0.15 per, the annual cost is miniscule: $450. This estimate can be increased by a factor of four to provide redundancy and mirroring, and another factor of four for the various encodings – and a final factor of 2 to compensate for any underestimate on the 1000 hours – and the result is still just $16k, which is about £10k.
[109] http://community.zdnet.co.uk/blog/0,1000000567,10012708o-2000331777b,00.htm#comment20101935
[110] http://crave.cnet.co.uk/software/0,39029471,49302215,00.htm

The total cost of storage for digital files is far higher than the cost of the media used and has to consider trends as well as regular migration.

Current options for cost-effective long-term storage at any scale all require automation through mass storage devices (disk servers and tape robots). The short obsolescence times for the components of these systems, including media, require the use of continual migration and active management. Traditional models of 'media on shelves' are unlikely to be cost effective, except for extreme cases where the need to access the material is very infrequent, e.g. once a year or less.

Our simple set of rules for calculating long-term storage costs are:

1. Take today's cost for commodity storage media (SATA hard disk drives or LTO4 tapes) and multiply by 10 to get the total annual cost of managed storage using that media.

2. Repeat step one if you want multiple copies, e.g. one copy on disk and one on tape.

3. Multiply the result of step two by 4 to give the TCO for a lifetime of storage

The cost of providing Internet access to audiovisual content can be many times higher than the long-term cost of storing that content.

For archive content that is regularly accessed, e.g. where there are more accesses per month than items of content, then the cost of providing Web-based access, including all the necessary infrastructure (e.g. transcoding, hosting, networking), even for 'browse' quality content, can be higher than the cost of preservation, including master copies. This has major consequences in terms of investment, business models and sustainability.

3. Preservation metadata models and extensions

Summary: We distinguish between preservation metadata and technical metadata.
- **Technical metadata**: ideally, the information that describes everything that needs to be known, in order to use (play, read, analyse, edit) the audiovisual content in a file.
- **Preservation metadata:** the information about a file that "supports activities intended to ensure the long-term usability of a digital resource"[111]

In the ideal case, the technical metadata would be so complete that no further information would be needed: rendering software could be constructed from the technical information. At present, for all audiovisual formats with the exception of uncompressed audio in .wav files, explicit technical metadata is not that complete, and so long-term preservation relies on naming of render/player software (at the very least). This situation is familiar to librarians: analogous to 'knowing someone who knows', preservation metadata names "an application which knows" how to render (or otherwise interpret) the contents of a file.

The main standard for preservation metadata is PREMIS[112]. In a typical digital preservation scenario, files of various kinds come to some sort of repository. There is some information about 'what they are' that comes with the files:
- implicit: an extension that gives weak and ambiguous identification of a 'file type'
- explicit: embedded information within the file. Sometimes (but rarely) the information will be formal metadata: expressed in XML with full schema and entity identifies (such as namespace designations). PB-Core[113] is an example of formal metadata, the vast SMPTE metadata dictionaries[114] constitute another, and MXF[115] is a file wrapper that could embed either (or both). More commonly, the information about what exactly the file is, and how it works, is embedded in the file in a way that is only interpretable by an application dedicated to that file type.

For example, while identifying information may be formally registered as a MIME[116] type (explicit, formal), the usual way to attempt to identify what a file 'really is', is through embedded "Identifying characters"[117] or "magic numbers"[118]. That information provides only the overall type identification. For audiovisual encoding, there is a range of parameters that also needs to be understood, and only an application that "understands" the file may be able to find and interpret those parameters. Having found them, such applications still may not make the information visible to the user, so that the user could write the parameter values into some set of formal metadata.

In such cases, PREMIS is used to hold information about how to play or otherwise use the file. The preservation metadata doesn't supply the missing technical knowledge – rather it names a player application ("I know someone who knows"), and that application is meant to be capable of understanding the embedded information.

---

[111] *Understanding PREMIS* by Priscilla Caplan for the Library of Congress
[112] http://www.loc.gov/standards/premis/
[113] http://www.pbcore.org/
[114] http://www.smpte-ra.org/mdd/
[115] http://tech.ebu.ch/docs/r/r121.pdf
[116] http://www.iana.org/assignments/media-types/
[117] http://filext.com/file-extension/MXF
[118] http://en.wikipedia.org/wiki/File_format#Magic_number

Identification and validation (based on syntax checking or magic numbers) is performed, for digital libraries, by well-established services: JHOVE[119], DROID[120]. DROID uses the PRONOM[121] database of file information. These services provide very limited support for audiovisual files (JHOVE doesn't support any video formats, PRONOM added limited information on MXF in October 2009), which is one of the purposes of PrestoPRIME: supporting audiovisual content within secure, formal repositories.

Preservation metadata for many file types has been compiled by the Library of Congress[122], which acts as a human-readable source of file format information, whereas JHOVE and DROID need machine-readable data. PRONOM data can be used by DROID, but also by humans. A very interesting line of research by David Tarrant of the University of Southampton has demonstrated a prototype system (P2)[123] of information on file formats compiled using *linked-data[124]*, which automatically amasses a broad range of data from the Internet. This approach eliminates the need for human compilation, and allows data to remain up-to-date.

It should be stated that current situation regarding file identification and validation is far from ideal. The ideal would be: all the information about the content pulled out and held in the file (or in a separate metadata file, but linked in an overall METS[125] structure) as formal metadata, to a recognised standard. Because of the wide array of file types and encoding types, many using proprietary technology and formats, and possibly having an effective lifetime of under a decade, the status of audiovisual files is likely to remain far from ideal.

The following sections will cover:
- a brief description of metadata standards for preservation metadata (PREMIS, and its use within a METS structure)
- the essential role of technical metadata, since this information is primary (as just discussed) – and how technical metadata is used in digital preservation
- the actual technical metadata that applies to audio and video
- how technical metadata is structured and stored, and how it could best be kept for preservation purposes

3.1.   Overview of preservation metadata models and standards

It is essential to distinguish between technical and preservation metadata. Technical metadata, if complete, would give the full definition of audiovisual content – where complete has an operational definition: the metadata gives all the information needed to play or otherwise use the content.

In contrast, PREMIS begins by excluding anything specific to a file type :

"Several categories of metadata are excluded as out of scope, including:
- *Format-specific metadata, i.e., metadata that pertains to only one file format or class of formats such as audio, video or vector graphics"[126]*

---

[119] JHOVE http://hul.harvard.edu/jhove/
[120] DROID http://sourceforge.net/projects/droid/
[121] PRONOM http://www.nationalarchives.gov.uk/PRONOM/Default.aspx
   also: http://www.nationalarchives.gov.uk/aboutapps/pronom/
[122] Library of Congress file format information: http://www.digitalpreservation.gov/formats/
[123] P2 http://de.scientificcommons.org/david_tarrant   http://de.scientificcommons.org/46567165
[124] Linked data: http://linkeddata.org/
[125] METS http://www.loc.gov/standards/mets/
[126] *Understanding PREMIS*, p4

This statement means that preservation metadata is not meant to describe content, technically or otherwise. Preservation metadata, at least as defined by PREMIS, acts at a higher level: it describes how content is to be used.

As the most obvious use of a file is to open it (look at it, render it), PREMIS supports defining not just names of players, but a whole IT environment (because a player could be embedded in a web browser that works under one operating system): *"Several semantic units are defined to record the environment of an object, that is, what hardware and software are required to render it and what dependencies there are on other objects"*.[127]

In addition, PREMIS can be used to define what really matters about the content. For instance, a document may be divided into pages. "Page" is a structural concept which may be vital for some documents (e.g. a scanned book that has been widely referenced using page numbers) – or it could be irrelevant (a document that was a long table or set of tables, where the pagination was a nuisance in the original).

The basic information supported by PREMIS is:
- a unique identifier for the object (type and value),
- fixity information such as a checksum (message digest) and the algorithm used to derive it,
- the size of the object,
- the format of the object, which can be specified directly or by linking to a format registry,
- the original name of the object,
- information about its creation,
- information about inhibitors,
- information about its significant properties,
- information about its environment,
- where and on what medium it is stored,
- digital signature information,
- relationships with other objects and other types of entities.

As can be seen from the above list, PREMIS is happy to refer to 'the format of an object'. Audiovisual files are not at all well-described by listing their 'format' – because that won't specify the encoding much less the detailed parameters. The complexities of audiovisual files will either be captured in significant properties' or in technical metadata. However PREMIS can be used to specify an environment – a player and operating system.

In summary, PREMIS supports holding information about applications that can play or otherwise interpret audiovisual contents. These applications, these environments, cannot be expected to last – hence the stress on 'significant properties' and *preservation actions* -- migration of 'file formats' to objects that work in new environments.

Technical metadata can do something better: it can fully describe the content, so that there is no dependence (in principle) on players and environments. This approach is perfectly simple for unencoded audio, relatively simple for unencoded video, and probably not at all possible for encoded signals.

---

[127] *Understanding PREMIS*, p9

Rather than draw too arbitrary a distinction between preservation metadata and technical metadata, including the standards suitable for containing or representing it, the important thing is the distinction between the scale and complexity of the metadata needed for compressed and uncompressed files. Because of the limited longevity of formats, players and media environments, the key thing is to capture and record as much of the metadata as possible that allows the environment, including players, to be reconstructed if necessary. Vastly more metadata is needed to describe how to do this for compressed formats than is required for uncompressed formats. Worse still these formats become obsolete much faster than uncompressed formats and hence there is the need to capture and manage this metadata relatively frequently if a succession of compressed formats is used for long-term preservation. This in turn impacts on the standards and structures needed, i.e. use of PREMIS, METS. The rest of this section looks in detail at exactly what this metadata is for audio and video content.

3.2.   The role of technical metadata in preservation of AV

The *technical metadata* required for effective preservation of audiovisual content is, essentially, the description of the technical structure of the stored content, down to the bit level. The information in the next section is a somewhat lengthy, because it attempts to describe the actual technical details of audio and video. Technical metadata is just a set of labels, and the information is meant to show what the labels actually mean.

There are two classes of audiovisual technical metadata:
- parameters of the audio signal and the video signal
- parameters of the file or wrapper that holds that signal

The first class, signal parameters, will always be needed so that the content can be properly played. The second class consists if information only needed for so long as the file or wrapper is needed, so in a sense this class holds *temporary preservation metadata* – information which can be dispensed with if the content is migrated to a new file/wrapper format (unless it is retained as a 'coding history' of the content).

For audio, the main signal parameters are:
- sampling rate
- the number of significant bits per sample (and how the bits are stored in bytes)
- the number of channels (separate audio signals e.g. two for stereo)

For video, the situation is more complex, but an overall standard for digital video (SDI[128]) specifies a whole set of parameters and their values:
- frames per second
- fields per frame
- lines per field
- samples per line
- bits per sample
- how the samples correspond to a colour space
- how sampling is distributed across the three dimensions that make up a colour signal

---

[128] *SMPTE 259M and variants* http://en.wikipedia.org/wiki/Serial_digital_interface

The second class is the technical description of file formats. This technical information is essential for playback (or any other use) of audiovisual content, but the technical information may not be fully contained in formal technical metadata. Rather the opposite is the general case: the formal technical metadata is commonly limited to stating the *file format,* or at most the encoding and the file wrapper format. This situation has arisen because file formats are complicated, and so metadata just attempts to give a name to the file type (and possibly encoding type) – so that an appropriate player can be found. The player holds the 'knowledge' of the encoding/file details.

Problems arise because of complexity. Many encodings can share a common wrapper, so that, for instance, a ".wav" file can contain many different ways to represent an audio signal, ranging from non-linear allocation of bits to samples to highly-compressed data. The situation for video is more complicated just because there are so many file and wrapper formats, as well as so many encoding possibilities. Further, two files made using the same encoder, and wrapped in the same type of wrapper, can still differ enormously. They could differ in their *compression parameters,* so that one MPEG-2 file (for instance) could be broadcast production quality at 50 Mb/s, while another could be completely unsuitable for professional editing, being much lower quality and not allowing edit at specific frames (because the compression averages across a *group of frames*[129]).

This whole situation is very unsatisfactory for long-term preservation, because key knowledge is 'embedded' in players, rather than being captured in formal metadata. Hence the survivability of the content is dependent upon the survivability of the players. Audiovisual content is not unique in this respect – text files are equally dependent upon software that can 'render' their contents. However there are ways to reduce the dependence:

- **better metadata**: successful rendering software can determine what kind of data it is dealing with, by reading and interpreting meta-information from the file; this information could in principle be 'pulled out' of the file and made explicit as formal technical metadata. The audiovisual industry would benefit from much more agreement on where and how to place metadata in proprietary file types, and on ensuring that *all* the decode parameters were part of that metadata.

- **simpler files**: most of the complexities of audiovisual content are to do with compression methods and interpretation of compressed data. Uncompressed audio is virtually self-describing (or needs no description, beyond: the following is a sequence of audio samples – just work out three parameters and it can play perfectly). Uncompressed video is more complicated but it also is 'just a sequence of samples'. Virtually nothing general can be said about the data in a compressed file, and attempting the playback of a compressed file of an unknown type could well prove futile.

3.3.   Overview of Audio and Video Technical Structure

Digital files are a sequence of bytes. However these bytes have to be interpreted, by an application that 'understands' a particular type of file (a file format, such as PDF or DOC or WAV). For the purposes of managing files in a formal digital library or repository, and maintaining the usefulness of the information contained in these files, it is considered

---

[129] http://en.wikipedia.org/wiki/Group_of_pictures

useful, perhaps essential, to have a formal description of the structure of the data in the file.

The following sections cover:
- the essentials of digital audio
- structure of WAV files
- metadata in WAV files, concentration on the Broadcast Wave File standard
- the essentials of digital video
- general structure of QuickTime (MOV), MXF and MPEG-2 files

3.3.1.   Audio

Structure

Uncompressed audio files should pose few problems for digital libraries and digital preservation. There is one pervasive file format (WAV), it has a simple structure (a header followed by a sequence of samples), the format has been in existence for nearly 20 years[130] and it shows no signs of needing replacement in the foreseeable future (meaning, in the next 20 years).

Audio is a signal, which for the purposes of this document is any continuous function representing measurements of a physical quantity. For sound recordings, the measurement is (typically) the air pressure at the face of a microphone. Digital audio is a sampled version of that continuous analogue measurement, and of course digital audio can be synthetically produced in any number of ways, without the necessity of measuring anything.

For all audio signals, with only the tiniest and most unlikely exceptions, the sampling is uniform in time. Hence the sampling can be defined by a single number, the sampling rate. The range of frequencies that can be captured by sampling is determined by the sampling rate. The theoretical limit of the frequency range is 0.5 times the sampling rate. Because filters and other engineering considerations are not perfect, in practical terms the frequency limit is about 0.4 times the sampling rate. For the most common kind of uncompressed digital audio, the signals on commercial CD recordings or their file-based equivalents, the sampling rate is 44.1 kHz, and the range of useful frequencies has an upper limit of about 18 kHz. "Normal hearing" – which tends to be defined by mass measurements on military recruits – has an upper limit of about 18 kHz, but hearing ability declines throughout adult life and by age 60 most people would have an upper limit to hearing of under 10 kHz.

There are many ways to use numbers to represent amplitude measurements arising from microphone signals. The simplest is a linear scale. Because hearing is sensitive to relative rather than absolute differences (small differences on small signals are as perceptible as are large difference in large signals), it is efficient to code audio on a logarithmic scale. Telephony has used 8-bit log scales since the 1970's. Large amounts of log-scale audio used to exist in WAV file formats, but the rapid decline in the prices for memory and bandwidth have led to WAV files now being widely understood to be synonymous with a linear representation of signal amplitude.

---

[130] http://www.digitalpreservation.gov/formats/fdd/fdd000001.shtml

Given that an audio file consists of time-uniform samples, and each sample is on a linear scale, there is only one remaining parameter: the number of bits per sample. The ratio of the largest to the largest sounds that can be represented by an audio signal is the same as the ratio between the largest number and the smallest (non-zero) number. This ratio represents the dynamic range of the captured (or synthetic) signal. As the smallest number is 1, the ratio simply equals the largest number – hence the dynamic ranges is determines by 'the number of bits' in the sampling. Human hearing has a range from detection threshold (fully acclimated in a very quiet environment) to destruction threshold (physical damage; ripping apart of the middle and inner ear) of 130 dB, or a linear range of 10,000,000,000,000 to 1 (in terms of sound power). As power increases according to the square of amplitude, that is 'only' a range of about 3,000,000 to 1. In binary numbers, about 22 bits would be needed to capture this entire range. It is no coincidence that the limit to performance of the very best audio sampling systems is also 22 bits. Most lab and studio setups struggle to get past 19 bits. CD quality audio is 16 bits, which gives a dynamic range of about 95 dB. As most quiet sitting rooms have a noise level above 30 dB, and as sound exposures above 90 dB in a workplace would require staff to wear hearing protection under UK, US and European legislation), a 95dB dynamic range is perfectly adequate for general purposes.

Given the overwhelming practical significance, for file-based information, of division of data into bytes, virtually all audio is now either 16 or 24 bit sampling. All the CD-standard audio is 16-bit, and anything that moves to any larger dynamic ranges automatically moves beyond two bytes to three bytes, and becomes labelled as 24-bit sampling – though the number of 'true bits' is rarely more than 19, and almost never more than 22. Synthetic audio can, of course, have an arbitrarily large dynamic range – but anyone producing such data would have great difficulty finding any way to present it to human listeners at its full range.

There is one complication with multi-byte numbers (such as 16 or 24-bit audio), which is that there are essentially two options: most significant byte first, or least significant byte first. This is the 'big-endian vs. little-endian' division in data representation http://en.wikipedia.org/wiki/Endianness that has been a complication for decades. RIFF files in general, and certainly WAV files, use little-endian notation[131]. The similar-but-different AIFF audio files (Apple) use big-endian notation.

WAV format

A WAV file is one type of the general class of RIFF file formats, agreed between IBM and Microsoft in the early 1990's[132]. It is similar to the AIFF format developed by Apple. It has a variable number of chunks of descriptive data at the start of the file, and a simple system for allowing software to identify the type and length of a chunk. Software can either interpret a chunk, or if it is unrecognised, the software can simply skip the chunk because it will know the length (number of bytes) to skip before the next chunk.

There is a mandatory "format chunk" chunk that gives the sampling rate, the number of significant bits per sample, and the number of channels (meaning whether there is one audio signal as in monophonic sound, two signals as is stereo, or more signals as in a multi-track studio recording), and the total number of bytes for one time sample across all

---

[131] http://www-mmsp.ece.mcgill.ca/Documents/AudioFormats/WAVE/WAVE.html
[132] http://en.wikipedia.org/wiki/Resource_Interchange_File_Format

the channels. The "bytes per sample" is just that last value divided by the number of channels.

The audio data is itself the final "data chunk", which again has an identifier and a length. The structure of the audio data is then simply a sequence of bytes, interpreted as follows:

```
<Ch1S1 Ch2S1 … ChnS1>
 <Ch1S2 Ch2S2 ... ChnS2>
 ...
 <Ch1Sm Ch2m ... ChnSm>
```

where Ch1 is track or audio signal 1, up to an arbitrary number n of tracks, and S1 is time sample one, up to an arbitrary number m of samples.

Each sample Sx is one or more bytes, and the numerical interpretation of the multiple bytes as a single integer is the one remaining complexity. For WAV files the bytes are ordered least-significant first = little-endian.

With regard to the future of wave files: the format is so simple, and so effective, that it is difficult to imagine why it would need to change (much). However should everything about WAV files be forgotten, the format is also so simple that once the metadata chunks are passed over, the remaining data can be identified and properly decoded very quickly. It only takes identifying bytes-per-sample and number of tracks, and a proper interpretation of the multiple bytes-per-sample as a single number, and then the audio is perfectly decodable. It would help to read the metadata, especially the sampling rate, but 'listenable' audio could be derived from a WAV file that had no intelligible metadata.

The following describes the audio in a CD-quality stereo file, with no extra RIFF chunks:

RIFF header	12 bytes
FMT chunk	24 bytes
Data chunk header	8 bytes
<Ch1S1–Ch2S1>	4 bytes; 1 stereo pair
<Ch1S2–Ch2S2>	4 bytes; 1 stereo pair
…	(continue as needed, at 44.1 such pairs per second of audio)
<Ch1Sm–Ch2m>	last 4 bytes in the file
(storage or file management system pads with zeroes to end of data block)	

Structure of data in a CD-quality stereo audio WAV file

Broadcast WAV

For the purposes of a general data structure, a broadcast WAV file[133] [134] is a WAV file, fully conforming to the WAV file specification. That specification allows an arbitrary number of header chunks, and the EBU has exploited that option to define some extra metadata that is of particular interest for broadcasting and for preservation.

The coding history metadata is relevant to preservation, as it can be used to list all technical processes that have been applied to the audio, from the time of its original

---

[133] http://en.wikipedia.org/wiki/Broadcast_Wave_Format
[134] http://tech.ebu.ch/docs/tech/tech3285.pdf

digitisation to the present. Other parts of the EBU chunk also are relevant to provenance, and general identification of the audio.

The EBU chunk is as follows:

```
broadcast_audio_extension typedef struct {
    DWORD ckID; /* (broadcastextension)ckID=bext. */
    DWORD ckSize; /* size of extension chunk */
    BYTE ckData[ckSize]; /* data of the chunk */
}
typedef struct broadcast_audio_extension {
    CHAR Description[256]; /* ASCII : «Description of the sound sequence»*/
    CHAR Originator[32]; /* ASCII : «Name of the originator»*/
    CHAR OriginatorReference[32]; /* ASCII : «Reference of the originator»*/
    CHAR OriginationDate[10]; /* ASCII : «yyyy:mm:dd» */
    CHAR OriginationTime[8]; /* ASCII : «hh:mm:ss» */

    DWORD TimeReferenceLow; /*First sample count since midnight, low word */
    DWORD TimeReferenceHigh; /*First sample count since midnight, high word */
    WORD Version; /* Version of the BWF; unsigned binary number */
    CHAR Reserved[254] ; /* Reserved for future use, set to "NULL" */
    CHAR CodingHistory[]; /* ASCII : « History coding » */
} BROADCAST_EXT
```

Description: ASCII string (maximum 256 characters) containing a free description of the sequence. To help applications which only display a short description it is recommended that a resume of the description is contained in the first 64 characters and the last 192 characters are use for details. If the length of the string is less than 256 characters the last one is followed by a null character. (00)

Originator: ASCII string (maximum 32 characters) containing the name of the originator/producer of the audio file. If the length of the string is less than 32 characters the field is ended by a null character.

OriginatorReference: ASCII string (maximum 32 characters) containing a non ambiguous reference allocated by the originating organisation. If the length of the string is less than 32 characters the field is ended by a null character.

OriginationDate: 10 ASCII characters containing the date of creation of the audio sequence. The format is « ',year',-,'month,'-',day,'» with 4 characters for the year and 2 characters per other item.
    Year is defined from 0000 to 9999
    Month is define from 1 to 12
    Day is defined from 1 to 28,29,30 or 31
The separator between the items can be anything but it is recommended to use one of the following characters: '-' hyphen '_' underscore ':' colon ' ' space '.' stop

OriginationTime: 8 ASCII characters containing the time of creation of the audio sequence.
The format is « 'hour,'-',minute,'-',second'» with 2 characters per item.
    Hour is defined from 0 to 23.
    Minute and second are defined from 0 to 59.
The separator between the items can be anything but it is recommended to
use one of the following characters: '-' hyphen '_' underscore ':' colon ' ' space '.' stop

TimeReference: This field contains the time-code of the sequence. It is a 64 bits value which contains the first sample count since midnight. The number of samples per second depends on the sample frequency which is defined in the field <nSamplesPerSec> from the <**format chunk**>.

Version: An unsigned binary number giving the version of the BWF, particularly the contents of the Reserved field. Initially this is set to zero.

Reserved: 254 bytes reserved for extension. If the Version field is set to zero, these 254 bytes must be set to a NULL value..

CodingHistory: Non restricted ASCII characters containing a collection of strings terminated by CR/LF. Each string contains a description of the coding process applied. Each new coding application is required to add a new string with the appropriate information. This information must contain the type of sound (PCM or MPEG) with its specific parameters :

PCM : mode (mono, stereo), size of the sample (8, 16 bits) and sample frequency,

MPEG : sample frequency, bit rate, layer (I or II) and the mode (mono, stereo, joint stereo or dual channel),

It is recommended that the manufacturers of the coders provide an ASCII string for use in the coding history.

Large Audio Files

Computing and file management systems have had a series of problems with limits on memory size and file size. The EBU and the AES have tried to produce well-defined options for dealing with audio that had to be broken up into multiple files, because of file-size limitations. As computer systems develop and maximum file sizes grow, these options become obsolete.

One change that is needed for some purposes is larger files. A RIFF file starts with 4 bytes spelling the identification code "RIFF" followed by four bytes giving the size (in bytes) of the whole file – meaning the file is limited to just under 2^{32} bytes, which is 4 GB. [Because some applications interpret the RIFF size as a signed rather than unsigned integer, they interpret a four byte integer as a number between -2^{31} and $+2^{31}$, which means that some applications only work on files below a 2GB limit.]

Four GB of stereo audio at CD quality is about six hours of audio – though for 24-bit samples at 96 kHz sampling that reduces to 2 hours. Some archives are now adopting 192 kHz sampling for some preservation purposes, meaning only one hour to reach the 4 GB limit.

The EBU has developed the RF64[135] alternative to the RIFF format to allow 64-bit size fields, a extra factor of 4x 10^9. This extension means that even 24-bit audio, at 192 kHz sampling, for 16 channels of surround sound, could go up to 250 million hours in duration.
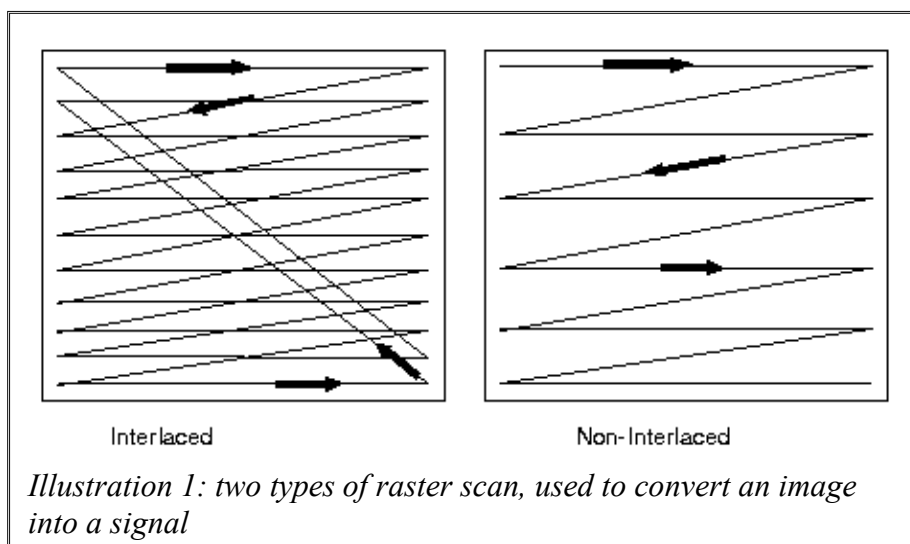
3.3.2.   Video

As with audio, digital video began as a sampled signal, with the difference that video is a signal representing a two-dimensional visual field, rather than the one dimension of air pressure at a microphone. A colour signal is more complex: three colour dimensions are measured across a two dimensional visual field.

Raster and interlaced

For television, but not for all images, the visual field is converted from two dimensions to, effectively, one by the mechanism of *raster scanning*: a two -dimensional area is divided into a number of lines, and then each line is divided into samples. The raster-scanning process divides a continuous two dimensional area into discrete lines, and the sampling process (digitisation) divides the lines into samples.

---

[135] [http://en.wikipedia.org/wiki/RF64]

Illustration 1: two types of raster scan, used to convert an image into a signal

This approach is not the same as dividing an area into 'little rectangles' – pixels – though when reproducing a digital video signal on a pixel-based display screen, a conversion into pixel specifications is required and can quite easily be performed. The major difference between a video representation and pixels is the issue of *interlace* – which affects both the temporal dimension of moving images, and the specification of colour information.

Interlace, like much else in audio and video technology, is a complication that arose because of technical limitations that no longer exist. Yet interlace must still be known and correctly interpreted, because interlaced signals sit on the video recordings in current archives – and are being converted to files that maintain interlaced data representations.

In order to eliminate flicker, television had to send images (*frames*) at a rate of about 40 per second, or more. Unfortunately the technology could only manage about 25 images per second. As more or less an expedient, half-images were created and transmitted at 50 half-images (*fields*) per second. For interlaced video, the odd-numbered lines are transmitted for one field, and then the even-numbered lines are transmitted for a second field. Thus the eye sees something like 50 'visual field refreshes' per second, and flicker is minimised.

For archiving, and for the purposes of data management and image integrity in the future, these fields cannot simply be converted into non-interlaced frames, for the simple reason that the television camera which produced the fields actually used a separate exposures for each field, and so the half-images are 20 msec apart in time (for UK/European systems with 50 fields per second). All future processing of interlaced data needs to take account of these time differences in order to maintain optimum image quality.

Colour

The other major complexity is colour. A black-and-white signal is has one dimension (brightness, luminance), but in television (and many other area of technology, e.g. printing) colour is treated as a total of three dimensions: brightness plus the two dimensions of a 'colour wheel' (chromaticity and saturation). Luminance plus these two colour dimensions amounts to a three-dimensional space represented by cylindrical coordinates. Equivalently the space can be represented in Cartesian coordinates, generally labelled YUV. An alternative coordinate system tied directly to the hardware of colour cameras and colour television tubes is RGB = red, green, blue.

A basic issue in all systems dealing with colour is finding and keeping a stable position in colour space. That whole topic is beyond the scope of this paper. The numbers in a video

file will represent (or can be converted to represent) values on the colour space dimensions, but it is a matter or convention and calibration across a whole industry – and not a matter of interpretation of the data in a particular file – that allows these values to give reasonably faithful colours.

The issue that is of importance in a given file, and that varies from file to file, is the way data is assigned to each of the colour dimensions. Video files can be expected to have data from the YUV dimensions, but in general there will be relatively more Y data (luminance; B&W information). Luminance Y defines the shape and gives the detail of the visual field, and the UV data just 'colours it in'. The eye is less sensitive to colour details, and so video files tend to sample the Y dimension two to four times as often as sampling the colour field.

If everything were equal, for every four samples of Y there would be four of U and V, for every field – represented by the notation 4:4:4. The overall data rate can be reduced, with little effect on picture quality, by having only two samples of colour for four of luminance, for every field, written as 4:2:2. The data can be further reduced by only sampling colour on one field of an interlaced frame, which effectively means 1 sample of colour per four of luminance, called (somewhat awkwardly) 4:2:0. The main types of colour video are 4:2:2 and 4:2:0.

The active region of a digital television image, sampled according to the standard for digital video which has been agreed since 1982 [CCIR recommendation 601; standardised for 10-bit data as *SMPTE 259M* = SDI = Serial Digital Interface], is 720 pixels by 576 lines for a frame rate of 25 Hz. Using 8 bits for each Y, U or V pixel, the uncompressed bit rates for 4:2:2 and 4:2:0 signals are therefore:

8-bit 4:2:2: 720 x 576 x 25 x 8 + 360 x 576 x 25 x (8 + 8) = 166 Mbit/s
8-bit 4:2:0: 720 x 576 x 25 x 8 + 360 x 288 x 25 x (8 + 8) = 124 Mbit/s

10-bit 4:2:2: 720 x 576 x 25 x 10+ 360 x 576 x 25 x (10 + 10) = 207 Mbit/s
10-bit 4:2:0: 720 x 576 x 25 x 10+ 360 x 576 x 25 x (10 + 10) = 155 Mbit/s

The above is for standard definition video (not the newer high definition formats), meaning a total of 625 vertical lines (including 'off-screen' invisible lines) in Europe and 525 in North America. The above is also for the 12:9 aspect ratio which was standard for video until the mid-1990's. Widescreen, at 16:9, would increase the above number by an extra 1/3.

The final parameter is bits per sample. There is no reason, beyond convenience, why Y, U and V (or their alternatives) should have an equal number of bits – but for simplicity the SDI standard calls for 10-bit sampling off all three dimensions: Y, U and V. High-end film scanning equipment (data cine[136]) commonly can produce 16-bit resolution video data through 'internal processing', though the dynamic range at the sensor is typically less than 60 dB (for the Spirit machine just cited, the figure is 55 dB), implying that there is no point in more than 10-bit sampling of the raw R, G, B signals.

Ten bits is an awkward number for the IT industry, as it is two big for a single byte and wastes 6 bits of a double byte. The SDI standard (for serial transmission) packs three 10-bit values into 32 bits (four bytes), 'wasting' only 2 bits of four bytes. For sequences of 4:2:2 data transmitted over SDI (or written to a file) as YUYV sequences, it is evident that 12 values (YUY-VYU-YVY-UYV) form the minimum group that holds a complete block of image data formed of packed 3x10-bit samples. So a basic low-level *data structure* for 10-

---

[136] ref Spirit http://www.dft-film.com/scanners/spirit_4k.php

bit video is not four bytes (as it would be for four 8-bit YUYV values), but 12 values occupying 16 bytes.

Other data in video files

Video is more complicated than audio in on other way: there is non-visual data that makes up an integral part of broadcast video files.

- **Audio:** Consumer analogue videotape had stereo sound from as early as VHS, but broadcast formats can have many more channels. File formats can support very large numbers of channels – a draft EBU standard covers up to 32 channels, and some file formats (e.g. MOV, MXF) can have an arbitrary number of audio channels.

- **Timecode:** because video signal from many sources would need to be mixed, and aligned with audio, CCIR 601 includes space in the video stream for time information, and this data has been included in professional video file formats.

- **Subtitles:** nearly all BBC output is now subtitled, and the text information (and metadata about its placement, colour and style) also need to be synchronised with video. CCIR 601 doesn't include direct provision for subtitle data, but MPEG and MXF formats (and others) do include subtitles – often multiple sets representing multiple languages.

- **Synch of audio with video:** audio to go with video can be in a separate file, and metadata and time code information can be used to make a stream or a new file with everything in synch – if all goes well. However it is generally seen as more secure to keep audio with the video: on the same digital videotape, and now in the same "video" file.

The basic method for combining non-video data, as just listed, is to consider the whole video stream as blocks, representing frames. Time code can then be specified with 'frame accuracy' (literal time codes can be placed in the file at every frame); subtitles information is an option that can exist at any frame – with a parameter giving the duration of the display of the subtitle; audio is broken in blocks with a duration equal to the frame interval (40 msec in Europe and other places with 50Hz systems).

"Unencoded" Video

The discussion so far has explained what's in a video file, but not given an actual structure. One issue is that most files don't consist just of the data so far described, but also use some form of coding to make the data more compact. The 'data model' for each compression type would be different, and would describe the layout and interpretation of the compression parameters – and then upon operation of a decoder a 'revealed' set of uncompressed data would have its own, further data model. There is probably very little of a general nature that can be said about compressed (encoded) video files, so this section will conclude with a couple of examples:

- two wrappers (file formats) for uncompressed video: MXF and QuickTime MOV

- one encoding: MPEG-2. MPEG-2 video can exist in 'its own' wrapper as an .mp2 file, or can exist in MXF, MOV or many other wrappers. Having already looked at MXF and MOV, the discussion of MPEG-2 will just outline the structure of the encoded data itself.

3.3.2.1. QuickTime MOV

QuickTime is an elaborate standard which is now in its tenth major version since its original release in the early 1990s. However there are three basic structures[137]:

- double fork: metadata and video data in separate 'forks' of a Macintosh file structure
- single fork: for simplicity, and compatibility with the PC world; metadata basically at the two ends of the file; metadata at the end is convenient for editing.
- fast start: a single fork file, but all metadata at the front so the file can start playing before it has all downloaded

Essentially a QuickTime file is also made of 'chunks' (called atoms), similar to the composition of a wave file. Each chunk begins with a 4-byte length, a 4-byte label (typically 'mdat' or 'moov') and then the data. The most basic single fork video has just two chunks: mdat with the data, followed by moov with the metadata. A 'fast start' file just puts the moov chunk first. Metadata within the moov atom would 'explain' about 4:2:2 vs. 4:2:0 sampling, and bits per sample, and everything else an application would need in order to correctly interpret the data in the mdat atom.
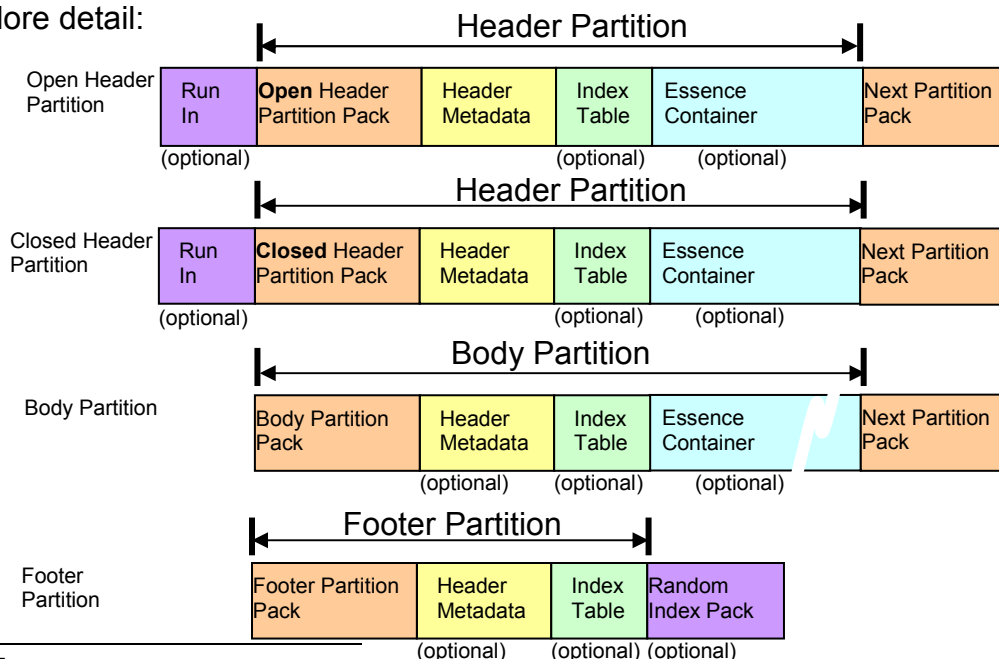
3.3.2.2. MXF

An overview and introduction to MXF was published in 2002 by the person with the major responsibility for producing the MXF standard, Bruce Devlin[138].

The following diagrams show the basic MXF file structure[139], consisting of a Header and a Footer; Body Partitions can be added.

---
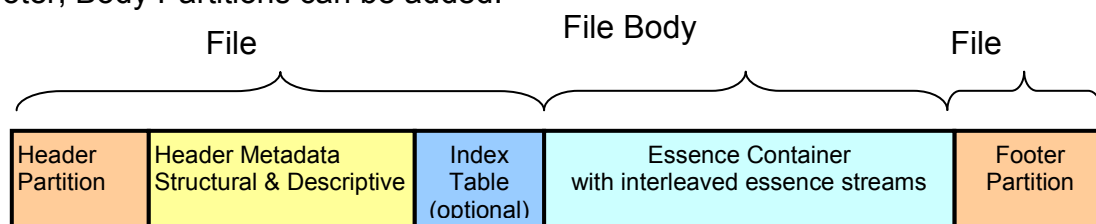
[137] QuickTime Toolkit: Basic movie playback and media types; Tim Monroe; Elsevier, p154
[138] http://www.ebu.ch/en/technical/trev/trev_291-devlin.pdf
[139] diagram from http://www.docstoc.com/docs/18285849/Metadata-Application-Review

Some aspects of MXF are unusual. A basic design requirement was support of streamed as well as file-based transport, hence the "Run In" block in the above diagram. Run-in is absolutely standard for videotape or audiotape: a portion of the signal that precedes the actual content, to allow equipment to 'get up to speed' and stabilize. An equivalent need exists for streamed media, even in the most modern technology: time for a media player and any associated technology to load and begin operating. MXF is a very general and far-sighted format, supporting a wide range of digital media requirements.

As a general format, MXF has a lot of options which in turn cause complications: which options should be used for a particular requirement, and will all the options be supported by all applications which work with MXF? The answer to the second question is 'no' – there have been issues of incompatible versions of MXF files and MXF applications[140].

As with MPEG encodings and other complex formats, MXF uses *operational profiles* (OP) to establish specific sets of options – specific ways of using the MXF format. For uncompressed video, the BBC has pioneered, and published, open source information and software: INGEX[141]. OP-1a (SMPTE 378M) is the profile for "a minimal simple MXF file" and is used by the BBC for storing uncompressed video in MXF. The BBC has been using MXF OP-1a for 8-bit 4:2:2 video originating as transfers from an obsolete digital videotape format, Panasonic D3. Current BBC work will extend to 10-bit data, so the process can also be used on standard SDI data arising from, for instance, the Sony Digibeta formats.

The "1a" part of the OP means that there is one item, and that it is contained in one package (container). The standard supports dividing an item into multiple containers, which could be useful to guard against risk of data loss during streaming, or to break large items into smaller pieces. However the straightforward implementation for MXF on one item is simple to use OP-1a and put that one item into one container.

OP-1a says nothing about encoding. The content could be HD or SD, it could be encoded as MPEG or DV, or it could be unencoded. The Header Metadata specifies the nature of the content.
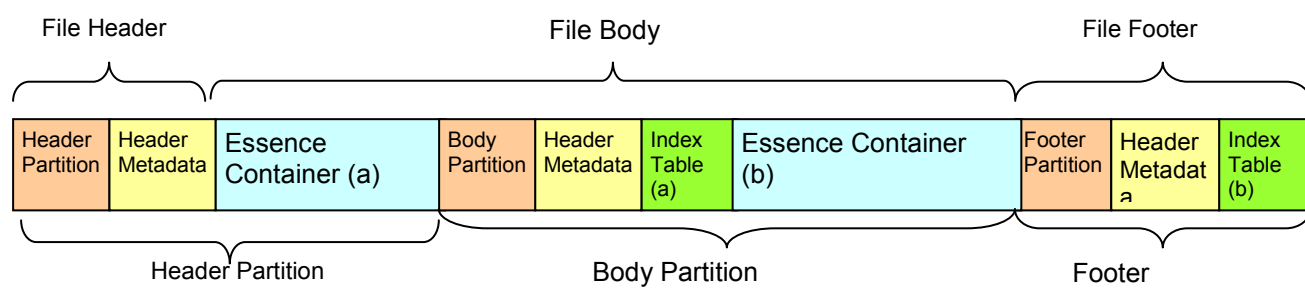
MXF was designed from the beginning to have an image (frame) as a basic unit. The Essence Container subdivides into frames, each coded as "one KLV packet". KLV is 'key length variable' coding, which is itself a SMPTE standard, and makes perfect sense for a world of dedicated digital links connecting things like studios and transmission areas. As such networks become replaced by 'standard' IP (internet protocol) networks, KLV itself may appear anachronistic. This summary cannot really explain the detail of MXF (which is

---

[140] http://downloads.bbc.co.uk/rd/pubs/whp/whp-pdf-files/WHP155.pdf

[141] http://ingex.sourceforge.net/

expressed in more than ten SMPTE standards), and for further information the basic SMPTE documentation should be consulted[142].

In outline, the BBC use of MXF for uncompressed video has the following approach[143]:

The OP-1A MXF file contains a sequence of content packages, one for each frame, containing video, audio and timecode, in the following formats:

- A video item of SMPTE 384M uncompressed 4:2:2 video at 8 bits per sample in UYVY format
- An audio item with 4 tracks of SMPTE 382M uncompressed PCM audio at 48kHz and 20 bits per sample
- A system item containing an array of SMPTE 12M timecodes representing the VITC and LTC timecodes read off the video tape

In addition to the essence, the OP-1a MXF file contains the following metadata in the MXF Header:

- Data items from the BBC Infax archive database record for the TV programme, identified by SMPTE Universal Labels registered under the BBC node of SMPTE metadata dictionary
- An array of D3 error records to facilitate future analysis, preservation and restoration
- An array of Photosensitive Epilepsy (PSE) failures if any
- The barcode of the LTO-3 data tape holding the MXF file

3.3.2.3. MPEG-2 compressed video data

The following diagram shows the layout of MPEG-2 data as a bitstream (sequence). How the data resides in files is slightly more complicated, but essentially (as with audio samples in a *wav* file), as a file the data shown would be preceded by identifying blocks. As discussed earlier, MPEG-2 data can be held in various file formats/wrappers (MPG, AVI, MOV, MXF…) – each with a different way of holding the metadata, but each eventually having a data sequence as shown here.

---

[142] SMPTE standards for MXF are summarised here: http://en.wikipedia.org/wiki/Material_Exchange_Format
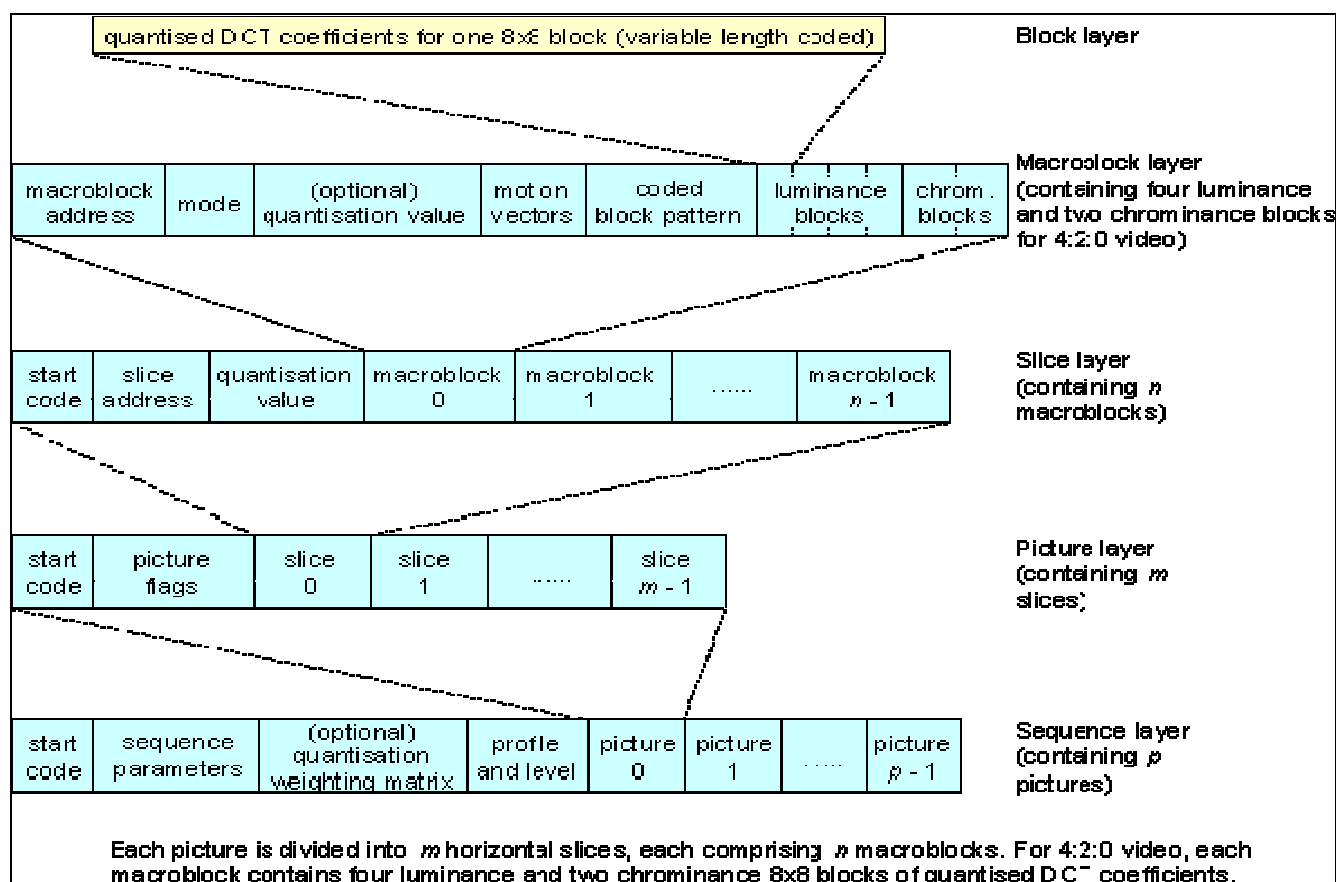[143] http://ingex.sourceforge.net/archive/

Illustration 2[144]: Outline of MPEG-2 video bitstream structure (shown bottom up).

3.4. Approaches to structuring and storing technical metadata

A basic problem with audiovisual files – though by no means unique to audiovisual content – is the proliferation of standards. In order to recommend a standard, or even a general approach to audiovisual metadata, this report starts with considering the function of the technical metadata, in hopes that such a consideration will help point to a solution.

Basic purpose: for PrestoPRIME, we focus on preservation: keeping audiovisual content usable, despite the complete change of all the technology used to hold that content. In the discussions above, we have argued that comprehensive technical metadata could give a complete description of audio and video signals held in a file (in the absence of compression), so that these signals could be recovered and used, with no dependence upon any existing player or rendering software.

We assume files will be preserved in a formal application: a digital library such as the Fedora/D-Space Duraspace[145], or a commercial system such as Rosetta[146], offered by the PrestoPRIME partner Ex Libris.

Getting files into a secure system: In a digital library approach, there is the following general flow of operations:

---

[144] Illustration reproduced with permission: www.bbc.co.uk/rd/pubs/papers/paper_14/paper_14.shtml

[145] http://duraspace.org/index.php

[146] Ex Libris Rosetta: http://www.exlibrisgroup.com/default.asp?catid={916AFF5B-CA4A-48FD-AD54-9AD2ADADEB88}&itemid={9B1F2C8A-3B03-459F-A2B4-4425A4D79689}

- **files enter:** action is taken to determine 'what they are'; DROID or JHOVE can be used to perform a very basic level of identification and validation
- **metadata is stripped out of files:** if metadata is going to be used for anything, it has to be accessed and made machine readable by the digital library system. If metadata is left embedded, then it will remain essentially *dark*[147] to all digital library functionality – and so essentially useless for digital preservation actions as undertaken by that digital library
- **metadata stored according to a standard:** the technical parameters embedded in the file will not, in general, come with labels. For example, there will be a number in an audio file representing sampling rate, but there will also be the following problems:
  - finding the parameter: where is it in the metadata?
  - knowing how it is coded: is it samples per second as a 16-bit unsigned integer, or what? Sampling rate is probably the simplest parameter of all. Broadcast Wave Format supplies a "time reference" which can be used to align audio with video, or to align multiple files of audio. It is coded, as shown above, as a 64-bit value defined as the number of samples since midnight. So its 'unit' is the sample, requiring knowledge of sampling rate in order to compute an 'actual' time. There is nothing in the Broadcast Wave File that explains any of this, or indeed even identifies where the 64-bits in questions reside. The digital library would need 'BWF knowledge' to find and interpret even the very simple BWF metadata. Ten years on from the development of BWF, we might wish for metadata in XML, according to a named schema with named entities (identified by a namespace). There will be none of that actually inside the BWF file, and so it is the task of the digital library system to create such a formal structure. Virtually all audiovisual files are in approximately the same state as for BWF: no schema, no labels, no namespace. The major exception is the class of proprietary file types which don't even have a published description (making BWF look easy).
  - knowing what formal schema to map the parameter to
  - knowing how to label the parameter as an element of that schema
  - performing the mapping, for all needed and available technical metadata, and writing the result in a standard fashion (probably as an XML file, in conformance with the chosen schema)
- **metadata and content integrated as an SIP using METS**
  - the metadata is now separate from the original file. The digital library needs to have a unit that is larger than the single file, in order to *package* multiple files. The common package for multi-file units in a digital library context is the METS(ref 125) standard, developed by the US Library of Congress and partners, under the overall structure of the Digital Library Federation.[148]
- **a catalogue of file content is constructed**
  - a digital library is of little help if the content of the library cannot be found. Hence an essential constituent is a catalogue, and this can be automatically constructed from the metadata stripped from the original file. The administrative and descriptive metadata is essential, though the catalogue may also include selected items of technical metadata. If the file doesn't have sufficient embedded metadata, then the formal solution is to create a

---

[147] dark metadata: http://www.w3.org/2005/Incubator/mmsem/XGR-vocabularies-20070724/ Section 3.3.3, final paragraph. The term appears to have originated with MXF, but the concept is general.
[148] Digital Library Federation http://www.diglib.org/

package (again, METS could be used) of 'submission information', so that the digital library can get its information from other files in the package.

Looking back at the steps just outlined, it should be apparent that the hard problem is pulling metadata out of an audiovisual file: finding it and properly interpreting it. Finding a metadata standard to hold the results of the extraction process is the easy part. As stated, there is no shortage of metadata standards!

A very useful contribution of PrestoPRIME will be tools for pulling all embedded and implicit metadata out of a small number of important audiovisual formats. Users of these tools can select their own metadata standard or standards for holding the result within their digital library or repository.

As an example of a comprehensive approach to use of various kinds of metadata within a formal digital library, the documentation of the Harvard Digital Repository Service[149] describes various standards that can be used in their repository. Their description does not explicitly state the degree to which the standards are supported by automatic tools.

In the absence of automation, files will need to be manually *acquisitioned* just as in a traditional book library. The importance of automation cannot be overstated: the manual approach works for a library that purchases a small number of items per week. With the conversion of various entire industries (including broadcasting) to a file-based approach, there will be thousands of new files a day for major enterprises – and even hundreds of files per day is too much for a manual process.

3.5. Key Findings and Conclusions

It is difficult to be generic about compressed (encoded) audio or video data: any compressed format is idiosyncratic by definition, and these idiosyncrasies need to be taken into account during preservation.

A file management or preservation process that needs knowledge of the structure of data within an encoded file will need the same information as in an application that plays or renders such a file. Given the large number of encodings for audiovisual data, plus the rate at which new ones develop and old ones become obsolete, it is an open question whether effort should be put into identifying a separate, formal data structure for any encoded audiovisual data. Instead, the effort could go into keeping the bitstream (binary representation; data sequence) intact, and keeping a valid and working player (codec) for each such file type.

Technical metadata can do something better (than can preservation metadata): it can fully describe audiovisual content, so that there is no dependence (in principle) on players and environments. This approach is perfectly simple for unencoded audio, relatively simple for unencoded video, and probably not at all possible for encoded signals.

The audiovisual industry would benefit from much more agreement on where and how to place metadata in proprietary file types, and on ensuring that *all* the decode parameters were part of that metadata.

---

[149] Harvard Digital Repository Service http://hul.harvard.edu/ois/digproj/metadata-standards.html#technical

Annex I : Calculating TCO for storage in perpetuity

T_∞ = total cost to store indefinitely
T_n = total cost to store for n years
T_1 = Current annual storage rate

α = number of years before annual cost of storage halves, i.e. cost half-life
β = rate at which annual cost falls

So, $\beta^\alpha = 0.5$

Cost for 1 year = T_1
Cost for 2 years = $T_1 + \beta T_1$
Cost for 3 years = $T_1 + \beta T_1 + \beta^2 T_1$

Cost for n years = $\quad T_n = T_1 \sum_{i=0}^{n} \beta^i \quad = \dfrac{1-\beta^{n+1}}{1-\beta}$

Cost in perpetuity = $\dfrac{1}{1-\beta} \quad = \dfrac{2^{1/\alpha}}{2^{1/\alpha}-1}$

Number of years for annual cost to halve	Total cost as a multiplier of year 1 cost
1	2
2	3.3
3	4.9
5	7.7
10	14.9

Annex II : Considerations on lossy compression and decoding

When a digital image is lossy encoded, e.g. a frame in a video, then information is being lost, or in other words artefacts are introduced when the image is decoded and displayed. If the decoder has some knowledge of how the digital image was encoded then it is potentially possible for this knowledge to be used to lessen the artefacts introduced by the encoder. This is not to say that the original image can be recreated, only that a new image can be created that is probably closer to the original image in some way than one created by a decoder that has no knowledge of how the encoder did its stuff. A trivial example is given later in this section to illustrate this concept.

Keeping knowledge of the encoder means this is available in the future to build better decoders. This of course might not be easy. If the encoder is open source software then no problem. If the encoder is proprietary hardware then there may be a real problem. The manufacturer isn't going to tell you how his encoder works, especially the details that might differentiate it from a competitor. Storing the physical encoder 'museum style' probably doesn't help either - you could loose the ability to run it and even if you can run it you could have a real problem trying to reverse engineer it to understand what it does in enough detail to provide the information needed to build a better decoder!

Example: Imagine a video sequence with the following frames 1,2,3,4,5,6... Imagine a lossy video encoder that does the following: (a) throw away every other frame from the video sequence so only the odd numbered frames remain: 1,3,5,7 (b) replicate these frames to create a new sequence 1,1,3,3,5,5,7,7... (c) compress the new sequence using ZIP or some other data compression tool. The result is a video sequence with the same number of frames as the original but half the storage space. Information has clearly been lost as every other frame is thrown away, so this is a 'lossy' encoding.

The 'standard' to be followed when decoding the sequence is simply to use ZIP and decompress. This gives the sequence of frames 1,1,3,3,5,5... that can then be played. Without any other knowledge of how the encoding has been done, this is the best that can be done. However, if there is knowledge that the sampling technique used by the encoder was to discard every other frame then this knowledge can be used to do a better job. For example, 'clever' stuff, e.g. motion compensated interpolation can be used to create the 'missing' frames. This would create a sequence 1,2',3,4',5,6',7,8'... The original frames 2,4,6... and the interpolated frames 2',4',6'... won't be identical but it is likely that 2' will be closer to 2 than simple reuse of frame 1.

In this way, knowledge of the encoder has allowed decoding to be done in a way that produces a result that is probably closer to the original than a naive decoder that does nothing more than follow the standard.

This is a hypothetical and very simple case. It does however have the main characteristics of conventional video codecs - first a lossy stage involving quantisation/sampling (DCT would be a real world example) followed by a lossless stage to achieve compression (e.g. entropy encoding). It is the first stage that introduces the losses. For complex and real-world encoding schemes, it may well be that once the image has been decoded with a less than optimal decoder then it will contain artefacts that wouldn't be present if a 'clever' decoder was used instead.

There are two important points here:

- If storing content in a lossy compressed format then as much knowledge as possible should also be stored on how the encoding was done.

- If decoding content, e.g. because it is now affordable to keep it uncompressed, then it is important that the best possible decoder is used to get the best possible representation of the original image.